Investigation of a novel *Ercc1* transcript predominant in skin

YVONNE A. SIMPSON

Thesis presented for the Degree of Doctor of Philosophy

College of Medicine and Veterinary Medicine

University of Edinburgh

2004

Declaration

The composition of this thesis and the work presented in it are my own, unless otherwise stated. The experiments were designed in collaboration with my supervisor, Prof. David W. Melton.

Yvonne A. Simpson
Edinburgh
December 31, 2004

Dedication

To my family

&

In loving memory of my grandparents

Mabel and David Gillespie

Table of Contents

List of Figures

List of Tables

Abstract

Name of Candidate: Yvonne Arlene Simpson Date: 31/12/04
Address: CRUK lab., MMC, Western General Hospital, University of Edinburgh
Degree: PhD
Title of Thesis: Investigation of a novel *Ercc1* transcript predominant in skin

The integrity of an organism's genome is constantly threatened due to the damaging effects of external and internal environmental factors. Additionally, the DNA molecule also has an inherent chemical instability. All these factors contribute to the formation of DNA lesions which, if not repaired, could develop into permanent mutations and malignancy. Fortunately, many DNA repair systems exist to prevent the accumulation of lesions. Of these, the nucleotide excision repair system (NER) is capable of repairing the greatest range of lesions, including those caused by UV irradiation, cross-linking agents and free radicals. Defects in the NER pathway disrupt DNA repair and one human disorder characterised by defective NER is *xeroderma pigmentosum* (XP). Exposure to UV is extremely dangerous for individuals with XP since their disorder predisposes them to a greatly increased risk of developing skin cancer. The gene *ERCC1*, which has an essential role in the NER pathway, has not been linked to a human disorder but mice in which the *Ercc1* gene has been inactivated are DNA repair deficient. These mice die, prior to weaning, of liver failure.

Earlier work in the Melton laboratory had demonstrated that there was tissue-specific *Ercc1* expression in the mouse and an additional, novel, *Ercc1* transcript was identified in the skin. The skin has a higher biological demand for repair of UV damaged DNA than other body tissues. As *Ercc1* is required for repair of UV damaged DNA the novel transcript may be involved in meeting the increased demand for repair in this tissue. The difference between the skin-specific and normal transcripts was localised to the 5′ end and is caused by differential initiation of transcription, with the skin-specific transcript initiating 5′ to the normal transcript's

promoter. This indicated the likely presence of a separate skin-specific promoter. To facilitate more detailed analysis of the upstream promoter region and the identification of sequences responsible for regulation of the skin-specific *Ercc1* expression pattern a number of *Ercc1* minigenes were constructed.

In this project the skin-specific pattern of expression was studied by transfecting *Ercc1* null cells with the minigenes. The *Ercc1* null cells used were murine embryonic fibroblasts, PF24^{+}#2D-4; Chinese hamster ovary cells, CHO43.3B; and murine keratinocytes, Ker. (-/-). UV survival studies were performed on transfected cells and RT-PCR was used to study the pattern of *Ercc1* expression. Transfection with the *Ercc1* minigenes corrected the UV sensitivity of the null cell lines. The *Ercc1* expression pattern was usually, but not always, appropriate for the cell type transfected.

Comparisons between the *Ercc1* transcription of *in vitro* cultures and *in vivo* tissues taken from various stages of murine development were made by means of northern blotting. Irradiation of transfected cells and primary cultures with UV-B, UV-C and visible light was performed in an attempt to identify possible sources of transcript induction. Changes in *Ercc1* transcription following irradiation of primary cultures were observed, but these changes did not conclusively prove that UV could induce *Ercc1* transcription. Serum starvation experiments were performed upon primary cultures to further study differences in *Ercc1* transcription *in vitro*. Release from serum starvation resulted in increased skin-specific transcript production and total *Ercc1* expression. The size of the skin-specific *Ercc1* transcript in keratinocytes was found to be larger than that of fibroblasts. Sequencing demonstrated that the size difference was in the CT repeat region of the gene. This region was deleted from one of the *Ercc1* minigenes mentioned and results indicated that it may be required for correct expression of the novel transcript.

Word count (excluding table of contents and references): 64,637

Acknowledgements

To begin with, I would like to thank my supervisor, David Melton, for giving me the opportunity to do this PhD. His guidance and support has been invaluable and I feel very fortunate to have been part of his research group. Thank you also to Ann-Marie Ritchie, Andrew Winter and Carolanne McEwan for teaching me the techniques required for the work presented in this thesis. I would also like to thank Jim Selfridge (as his work laid the foundations for mine), Rachel Berry (for her kind offers to develop film when things got really busy), Kate Britton (for tracking down alternative sources of lab. supplies at short notice) and Joyce Begbie (for last minute autoclaving and solution top-ups for the out-of-hours work). The author of the ImageJ software, Wayne Rasband, also deserves special thanks for his advice upon the use of his software for analysis of gel bands. My work was funded by the M.R.C. and the Orkney Educational Trust kindly awarded a grant towards the costs of thesis printing.

Of course, I am grateful to all the members of the Melton Lab. for their friendship and encouragement (including those already mentioned). It has been a pleasure to get to know them all and I hope that the friendships formed during my time here will continue to grow for many years to come. There is a Turkish saying that I would like to mention before I thank everyone and it roughly translates as so - just because I praise her does not mean you are not worthy of equal praise. So, if I forget to mention anyone (or any favours) be assured that it is a result of my writing-induced memory-loss and does not mean that I'm not grateful. In addition to those already mentioned, I would like to thank the following lab. members (past and present), ladies first: Carolanne, Claire, Jennifer, Joanne, Kristina, Lesley, Lynn, Marion, Marina, Mary, Niki, Roberta and Sarah. Now, thanks to the gentlemen: Charles, Martin, Michael, Neil, Paul, Scott and Yan. I would also like thank some others who work in the Molecular Medicine Centre: Alan (for cheery welcomes upon arrival in the building each morning) and all my tea-break friends from the other research groups.

I would like to acknowledge the importance of the following people to my time in Edinburgh: Tessa Spencer, for being a lovely flatmate and friend; Dr. Jon Turner, for the invitation to join the postgrad science communication team; friends from Technopolis at the Edinburgh Science Festival and Madeleine Shepherd, for being a great neighbour and science festival partner.

I would also like to thank the following inspirational teachers/lecturers who played their part in the educational process: Mr. James Curran, Mrs. K. Holmes, Mr. I. Ballantine, Prof. P. Wilmer, Prof. Slater, Dr. Martin Milner, Prof. Cole-Hamilton, Dr. John McLachlan, Prof. J. Turner and Dr. Cobb.

Finally, I would like to offer my deepest gratitude to my family and to Alasdair for their love and support.

Abbreviations

~	approximate
λ	wavelength
μ	micro
6-4PP	(6-4) pyrimidine pyrimidone photoproduct
A	adenosine
AB sites	abasic sites
AP sites	apurinic/apyrimidinic sites
BER	base excision repair
bp	base pair(s)
C	cytosine
C-	carboxy-terminal
cDNA	DNA complementary to RNA
CHO cells	Chinese hamster ovary cells
CPD	cyclobutane pyrimidine dimer
CS	Cockayne's syndrome

dNTPs	deoxynucleoside triphosphate
DMEM	Dulbecco's modified Eagle Medium
DMSO	dimethylsulphoxide
DNA	deoxyribonucleic acid
DNA pol	DNA polymerase
dRP	deoxyribose-phosphate moiety
DSB	double-strand break
EDTA	ethylenediaminetetraacetic acid
EGF	epidermal growth factor
ES	embryonic stem
ERCC	excision repair cross-complementing
EtBr	ethidium bromide
FCS	foetal calf serum
g	gram
G	guanosine
GGR	global genome repair Transcription coupled repair (TCR).

GMEM	Glasgow Modified Eagle Medium
HPRT	hypoxanthine phosphoribosyltransferase
HR	homologous recombination
IR	ionising radiation
J	joule(s)
kb	kilobase
l	litre(s)
L	lower (normal) *Ercc1* transcript
LB	Luria broth
m	messenger
m	metre
M	molar
MMR	mismatch repair
MOPS	3-[*N*-morpholine] propane sulphonic acid
N-	amino terminal
NER	nucleotide excision repair

NHEJ	nonhomologous end joining
O^6-MGT	O^6-methylguanine transferase enzyme
OD	optical density
OLB	Oligo labelling buffer
PBS	phosphate buffered saline
PCNA	proliferating cell nuclear antigen
PCR	polymerase chain reaction
PGK	phosphoglycerate kinase
PR	photoreactivation
r	ribosomal
RNA	ribonucleic acid
SSA	single-strand annealing
ssDNA	single stranded DNA
T	thymidine
TB	terrific broth
TCA	trichloroacetic acid

TCR	transcription coupled repair
TLS	translesion synthesis
TTD	trichothiodystrophy
U	Upper (skin-specific) *Ercc1* transcript
UTR	untranslated region
UV	ultra violet
V	volts
wt	wild type
XP	*xeroderma pigmentosum*

Chapter 1

Introduction

1.1 Foreword

DNA is the fundamental substance of which genes are composed and is often referred to as the 'genetic blueprint' of life. As a chemical data store, DNA has to be stable yet some flexibility is required. Occasionally DNA mutations may occur that confer advantages to the host organism and, in this way, the instability of the DNA molecule is the substrate of evolutionary progression. There is a balance between the data of the genome being stored as an incorruptible medium yet also allowing for change.

Rarely a mutation may indeed be advantageous to an organism but most DNA mutations have either no effect whatsoever or are deleterious in nature. When lesions are not repaired, DNA damage can lead to the formation of permanent mutations, resulting in genome instability, carcinogenesis and ageing. Because of the likelihood of life-threatening outcomes such as carcinogenesis, it is of utmost importance to an organism that mutations are repaired quickly and for irreparable DNA damage to be removed via cellular apoptosis. The likelihood of mutations being deleterious has exerted a selection pressure over time and a number of DNA repair mechanisms have evolved. The complexity and number of an organism's DNA repair mechanisms is dependent upon the size of the organism's genome.

DNA mutations can occur spontaneously as a result of the inherent instability of the chemical bonds of the nucleotides at normal physiological temperature and pH. The genome is a dynamic molecule, undergoing replication and recombination; mutations can arise due to errors in these processes. Further mutation can occur as a result of incorrect repair of DNA damage and the chemical reactions taking place in the cellular environment, such as oxidation, hydrolysis and non-enzymatic methylation, can result in DNA mutations. Chemicals entering an organism's system from the environment, be they naturally occurring or man-made, can result in DNA damage if they are capable of altering cellular chemistry or directly interacting with DNA. In the natural environment, ultra violet (UV) radiation from the Sun is a major source of DNA mutation in many organisms.

The study of DNA repair mechanisms began in prokaryotic systems. However, the greater complexity and size of the eukaryotic genome meant that the research in prokaryotic systems could lead to greater understanding of eukaryotic DNA repair in some cases but was not directly applicable. For example, the multicellular nature of higher eukaryotes resulted in the evolution of pathways that prevent cells with damaged DNA passing through the cycle of division before repair has taken place, reducing the risk of carcinogenesis. Cancer is one of the leading causes of death in the United Kingdom, so research of mammalian DNA repair mechanisms is vital to the improvement of cancer prevention and treatment.

1.2 DNA damage and causal factors

DNA damage may be classified in a number of ways. If damage is classified by source then DNA damage is either from an exogenous source (caused by external, environmental factors) or an endogenous source (caused by internal factors). When the 'environment' is that of the organism rather than the DNA molecule then cellular sources of DNA damage are endogenous. DNA damage may also be classified by whether the damage occurs spontaneously or is induced; spontaneous damage arises due to the inherent instability of the DNA molecule while damage caused by other factors is induced. However, the classification of damage is complicated by the fact that different sources of DNA damage can result in identical chemical changes to the DNA. Because of the large number of DNA damaging agents, both physical and chemical, a table (Table 1.1) of those most commonly encountered is provided along with the type of DNA damage resulting from exposure to such agents (reviewed by Friedberg, Walker and Siede, 1995). There are two categories into which DNA mutations can be divided and the categories are essentially opposites; the first is when the nucleotide is structurally intact but in the wrong structural context, the second is when the nucleotide is in the correct part of the DNA sequence but is chemically abnormal. Only the second category is classified as DNA damage or lesion and examples are fragmented, cross-linked or modified nucleotides (Sancar, 1995).

Table 1.1 Types of DNA damage induced by chemical and physical agents

Table shows a selection of DNA damaging agents, the source of such agents and the types of DNA damage arising from exposure.

DNA damaging agent	Source	Resulting DNA damage
Ultra violet radiation (UV) (wavelength dependent)	Physical environment	Cyclobutane pyrimidine dimer (CPD) (6-4) pyrimidine pyrimidone photoproduct (6-4PP) Interstrand cross-linking (minor product) Strand breaks
Ionising radiation (IR)	Physical environment	Strand breaks Base and sugar damage
Reactive Oxygen species e.g. superoxide anion radicals, H_2O_2.	Normal cellular metabolism Minor by-products of exposure to UV, IR	Strand breaks Oxidised bases and sugars Abasic sites (AB sites)
Alkylating agents	Environmental	Addition of bulky adducts e.g. methyl groups to nucleophilic centres of bases and phosphodiester backbone
Cross-linking agents (special category of alkylating agents) e.g. mitomycin C	Environment (chemical)	Intrastrand cross-links due to nucleotide modification Interstrand cross links if alkylating agent is bifunctional

1.3 Cellular responses to DNA damage

Because the genome of a cell encodes the genes required for production of protein in the cell, the viability of a cell is directly related to the condition of its DNA. Damage to the genome can lead to alterations in the proteins encoded by mutated genes, leading to loss of protein function. DNA lesions may also result in incorrect temporal or spatial gene expression, complete loss of gene expression or disruption of signalling pathways. Cells that accumulate mutations affecting their growth rate and responses to cell division check-points may divide in an uncontrolled manner, giving rise to tumour formation in the body. Because of the possible negative outcomes arising from DNA mutation and lesion formation it is crucial that a cell responds to DNA damage.

There are three possible ways in which a living cell can respond to DNA damage; *damage reversal*, *excision* or *tolerance*. The DNA damage is only repaired in the processes of reversal and excision. Repair of the damage occurs where the abnormal change in DNA chemistry or sequence is reversed by means of direct repair or excision of damage and consequent synthesis of replacement nucleotides. Where cellular tolerance of DNA damage takes place the damage is not repaired but is instead bypassed during the processes of DNA synthesis and replication. The complexity of DNA repair mechanisms is related to the size of the genome they are associated with as well as the nature of DNA mutation they repair and the context in which the DNA mutation is encountered in, i.e. the urgency of the repair.

1.4 DNA damage tolerance mechanisms

Unrepaired DNA damage can block replication forks, as bases cannot be inserted opposite the damage site by the normal replicative polymerases. In the following models of DNA damage tolerance mechanisms, referred to as postreplication repair, the DNA damage on the template strand is not repaired but bypassed during DNA synthesis. There are three models for how this can take place, template strand switching, recombination and translesion synthesis (reviewed in Friedberg *et al.*, 1995). In template strand switching the DNA polymerase (DNA pol) switches to the template daughter strand of the complementary DNA strand to bypass the lesion by

copying the newly synthesised daughter strand and then switches back. In the recombinational repair model, the sister chromatid provides the template for bypassing the lesion. In eukaryotes, the primary mechanism for postreplication repair is thought to be translesion synthesis (TLS) (Lehmann, 2000; Kunz *et al.*, 2000). In the TLS model, damage-bypass DNA polymerases can either perform mutagenic or non-mutagenic TLS. The human disorder *xeroderma pigmentosum* variant (XP-V) is mentioned in the section on human inherited disorders (1.9.1). The disorder in XP-V is caused by mutation in the *XPV* gene, which encodes the bypass DNA polymerase η (Masutani *et al.*, 1999). The DNA-damage bypass polymerase η performs non-mutagenic TLS while DNA pol ζ is a mutagenic bypass polymerase. In non-mutagenic bypass the DNA damage is bypassed and synthesis resumes downstream of the DNA damage resulting in a gap in the daughter strand that is repaired using the undamaged sister strand as a template. The repair of the gap has been referred to as daughter-strand gap repair to emphasise the fact that repair of the original DNA damage does not take place but repair of the resultant gap in the daughter strand does (Hanawalt, *et al.*, 1979). In mutagenic TLS the damage bypass polymerase inserts bases opposite the lesion.

1.5 Direct Reversal of DNA damage

The most simple way to repair damaged DNA is with an enzyme consisting of a single protein that is capable of directly reversing the DNA damage. Such a mechanism is likely to carry less risk of introducing DNA damage than a complex multi-step, multi-protein repair mechanism but is not capable of recognising and repairing as wide a variety of DNA damage.

1.5.1 Photoreactivation of DNA

An example of a DNA damage direct reversal reaction is a mechanism called photoreactivation. Photoreactivation (PR) reverses the most common type of DNA damage that occurs following UV irradiation, the cyclobutane pyrimidine dimer (Cook, 1970). As shown in table 1.1, irradiation of DNA with UV of a **wavelength**

near the absorption maximum of the molecule leads to the formation of cyclobutane pyrimidine dimers and 6-4 pyrimidine-pyrimidones. Such base damage causes disruption of DNA transcription and replication because dimerization of the bases produces distortion in the DNA helix. (Moore and Strauss, 1979; Setlow *et al.*, 1963; Swenson and Setlow, 1966; Villani *et al.*, 1978). The photoreactivation repair mechanism is present in many prokaryotes and eukaryotes but it is not found above the evolutionary level of marsupials in mammals (Li *et al.*, 1993; Yasui et al., 1988), although light-dependent loss of pyrimidine dimers from DNA has been observed in cultured cells (Sutherland and Oliver, 1976). The most widely studied DNA photolyase is that of *E. coli* and under normal growth conditions there are only 10 to 20 molecules of DNA photolyase in each *E. coli.* (Harm, 1970). In *Escherichia coli* (*E. coli)*, monomerization of the pyrimidine dimers by PR takes place via the action of the DNA photolyase enzyme; the catalytic activity of the enzyme depends upon absorption of light by its two chromophore co-factors. The reaction mechanism involves excitation of one chromophore by absorption of a single photon followed by transfer of excitation energy to the second chromophore. Excitation of the second chromophore results in electron transfer to the cyclobutane pyrimidine dimer and this electron transfer makes the dimer revert to a monomeric state. This monomerization process is referred to as cycloreversion (Sancar, 1994 and 1996).

1.5.2 Removal of alkylation damage

Examples of DNA alkylating agents are methylnitro-nitrosoguanidine (MNNG) and methylnitrosourea (MNU). Alkylation of DNA bases results in the formation of O^6-alkylguanine and O^4-alkylthymine; the alkylating agents are mutagenic because the formation of O^6-alkylguanine and O^4-alkylthymine lesions can result in mismatching of bases during semi-conservative DNA synthesis (Swann, 1990). Methyltransferase enzymes, that remove the methyl group from the base, directly reverse the alkylation. For example, in repair of O^6-methylguanine lesions the methyl group attached to the O^6 of the alkylguanine is transferred to a cysteine residue of the O^6-methylguanine transferase enzyme (O^6-MGT). The O^6-MGT enzyme is classified as a 'suicide enzyme' because it is permanently inactivated by transfer of the methyl group to the cysteine residue (catalysts are, by definition, not consumed by the reactions they catalyse) (Sancar, 1996; Teo *et al.*, 1984).

1.5.3 Single strand break repair

A final example of direct reversal of DNA damage is the repair of single strand breaks which result from hydrolysis of the phosphodiester bonds of the DNA sugar-phosphate backbone. The breaks can only be directly repaired if the damage results in a clean break with juxtaposed 3′ OH and 5′ P termini. Where this is the case, DNA ligase catalyses the reconnection of the strand ends. This process has only been observed in a limited number of organisms and damage by ionising radiation more commonly results in extensive damage to the bases and sugars at the break-point, necessitating more complicated repair prior to ligation (Sancar, 1996).

1.6 Double Strand Break repair

Double-strand breaks (DSBs) in DNA are a serious threat to genome integrity. This is because if they are not repaired then the part of the chromosome that has broken off (and is no longer attached to the centromere) is lost during cellular mitosis. For this reason, a single DSB can be lethal to a cell. DSBs may be formed by exposure to a number of DNA-damaging agents or during normal cellular processes. Environmental causes of DSBs include ionising radiation, free radicals, endonuclease activity (part of DNA repair process) and mechanical stress. DSBs can also occur due to a number of recombinational events that are part of normal cellular processes such as meiotic or V(D)J recombination (lymphoid site-specific recombination). If DSBs are not repaired or missrepaired then mutations or chromosome rearrangements may take place and eventually lead to cell death. (Reviewed by Wood, 1996).

1.6.1 Non-homologous end joining

Nonhomologous end joining (NHEJ) is the major pathway for repair of DSBs in multicellular eukaryotes, the other being homologous recombination (HR). In mammals the main pathway for DSB repair is NHEJ while in lower eukaryotes, such as yeast, the main pathway is HR. Depending upon the extent of damage at the site of a DSB, there may be a requirement for additional repair before ligation of the DSB can take place. The term 'nonhomologous' is a reference to the lack of

extended regions of homology between the DNA strands at the repair site. Unlike the situation in HR, regions of homology at the site of the DSB are not required for NHEJ to take place, so blunt ends as well as protruding complementary and non-complementary protruding ends can be repaired by NHEJ. Where regions of microhomology are available, in overhanging single strands at the double strand break, they may be utilised by NHEJ. Even a single complementary base pair between overhanging ends of DNA is sufficient for NHEJ to take place, although strands are aligned to make the maximum number of base pairs. NHEJ is also required to rejoin site-specific DSBs introduced during V(D)J recombination (Roth and Wilson 1986). Mammalian cells deficient in NHEJ are hypersensitive to ionising radiation (which causes DSBs) and are immunodeficient because of their inability to complete V(D)J recombination (Baumann *et al.*, 1998, Pastwa *et al.*, 2003).

The main components required for NHEJ in eukaryotes have been identified via *in vitro* studies. Each of the proteins involved in NHEJ either binds the ends of linear DNA directly or through protein associations. X-ray irradiation leads to DSB formation and study of X-ray hypersensitive mouse cell lines led to discovery of the XRCC genes. The following proteins are required for NHEJ; DNA Ku protein, DNA-PK(cs), XRCC4 protein and DNA ligase IV. Studies on yeast mutants have shown that a complex of RAD50, Mre11/Xrs2 and a family of Sir proteins (Sir2p, Sir3p and Sir4p) are also required for NHEJ. The DNA Ku protein is a heterodimer of Ku70 and Ku80 and is responsible for recruiting DNA-PK(cs) to the site of the DSB. DNA-PK(cs) is the catalytic subunit of DNA protein kinase and the whole DNA protein kinase complex is composed of the DNA Ku heterodimer plus this catalytic subunit. The DNA protein kinase regulates kinase activity through its association with the DNA. XRCC4 and DNA ligase IV associate to form a complex that is able to ligate one strand of the DNA with minimal base pairing to the antiparallel strand. Murine 'knockouts' lacking DNA ligase IV and XRCC4 are embryonic lethal because of their inability to perform NHEJ (Lieber, 1999). Depending upon whether repair requires removal of overhanging DNA (exonucleolytic degradation) at the site of the DSB or resynthesis of a missing region

(gap filling), a nuclease or a polymerase is required (Ma *et al.*, 2004; Pastwa and Blasiak, 2003; Jeggo, 1998; Lewis and Resnick, 2000). The single DNA strands cannot be individually ligated (no single strand DNA ligases have been identified) so an alignment protein is probably required to hold the strands in place while any necessary repairs are performed prior to ligation of both strands with DNA ligase IV. DNA polymerase has been proposed as the alignment protein (King *et al*, 1994). NHEJ commonly results in small deletions that extend in both directions from the DSB site and end within the region of microhomology (misjoining). 'Misjoining' typically occurs in regions of short repeats around the DSB. When the DNA is aligned according to microhomology there may be overhanging DNA sequences that extend beyond the end of the homologous region and these overhanging ends need to be removed by the action of a nuclease. The polarity of the DNA strand determines the type of nuclease required; 5′ nucleases FEN1 and EXO1 and the 3′ nuclease MRE11 have been proposed as likely candidates (Wu *et al.*, 1999; Lieber, 1999).

1.6.2 Classical homologous recombination

DSB repair by means of homologous recombination, in contrast to NHEJ, results in faithful restoration of the sequence at the break point. Because extensive damage can be present at a DSB point, repair by homologous recombination requires a non-damaged template (the homologous chromosome to the one containing the damaged DNA) for restoration of any damaged or missing sequence. Because of the requirement for interaction between DNA strands from the DSB and those of the template this is a more biologically complex repair process than NHEJ. The 5′ strand ends at the DSB are digested by nuclease activity to yield 3′ single-stranded overhanging tails that may be up to 1kb in length. In *Saccharomyces. cerevisiae* (*S. cerevisiae*), the *RAD52* epistasis group of genes encode a protein that binds to these 3′ ssDNA ends and this complex then invades the homologous intact duplex DNA sequences of the undamaged chromosome that acts as the template for repair. This strand invasion initiates strand exchange which leads to formation of joint molecules between the damaged invading strands and the undamaged template strands called Holliday junctions. The invading strands initiate repair synthesis and once the

missing DNA has been replaced the Holliday junctions are resolved by a resolvase that cleaves and separates the duplexes.

1.6.3 Single strand annealing

Single-strand annealing (SSA) is a special type of homologous recombination but, unlike classical homologous recombination, repair of DSBs by SSA is a non-conservative mechanism (reviewed by Pastink *et al.*, 2001). Although SSA does not involve the formation of Holliday junctions the process does depend upon regions of homology (direct repeats flanking both sides of the DSB) for alignment of the damaged strands prior to repair and ligation. The 5′ strand end digestion that takes place prior to homologous recombination leads to production of 3′ single stranded overhangs. These overhangs can be used to prime homologous recombination if a region of homology in the corresponding intact chromosome is located. However, if short regions of homology between the 3′ overhangs can be found then annealing can take place, resulting in repair by the SSA mechanism. The regions of homology may be as small as a 30bp repeat and SSA can take place with repeats as distant as 15kb from the DSB (Sugawara *et al.*, 2000). Although the factors required for SSA have been identified in yeast (only in *S. cerevisiae*) the genes involved in SSA in higher eukaryotes have yet to be determined.

Once annealing of homologous regions in the single stranded 3′ ends has taken place the non-homologous regions of the 3′ ends must be removed prior to ligation. If the 3′ non-homologous overhanging tails are shorter than 30nt then they are removed by a mechanism dependent upon the 3′ to 5′ proofreading ability of DNA polymerase δ (Paques and Haber, 1997). Tails longer than 30nt are removed by a different mechanism, requiring the use of a number of proteins. The endonuclase activity of the RAD1/RAD10 complex (homologous to the mammalian ERCC1/XPF complex) is necessary (Ivanov and Haber, 1995). RAD52 and RAD59 as well as the RAD50/Xrs2/Mre11 complex are also involved (Ivanov *et al.*, 1996; Sugawara *et al.*, 2000). RAD52 and RAD59 both anneal complementary DNA strands. The proteins in the RAD50/Xrs2/Mre11 complex all have endo- and exonuclease activity (Pastwa and Blasiak, 2003). The mismatch repair protein complex, MSH2/MSH3, is

necessary for SSA if the length of the annealed region is less than 1.2kb. This complex initiates incision of the DNA by RAD1/RAD10. Following annealing, RAD1/RAD10 make an incision at the junction between the annealed and non-annealed regions of the 3′ tails. Therefore, the role of MSH2/MSH3 when the annealed region is small may be the stabilisation of intermediate structures prior to nuclease activity and/or recognition of single-stranded loops at regions of non-homology (Sugawara *et al.*, 1997). The Srs2 protein is also required for SSA and this protein has a 5′ to 3′ helicase activity (Paques and Haber, 1997). Once the non-homologous 3′ ends have been removed ligation takes place. Because the region between the two repeats used for annealing is deleted, repair of DSBs by SSA is non-conservative.

1.6.4 Interstrand DNA Cross-links

Interaction between DNA and a number of chemical agents can lead to formation of cross-links. DNA cross-links can either form within or between DNA strands, respectively termed intrastrand and interstrand cross-links. Of the two types of cross-link it is the interstrand cross-link that is particularly damaging as it prevents DNA strand separation, blocking DNA replication and transcription. A single unrepaired DNA cross-link may result in cell death. Molecules that react with DNA to make a cross-link form a molecular bridge between the DNA molecule(s) and so such chemicals either have two reactive groups capable of covalently bonding to DNA or a dimer forms between two molecules that each have a reactive group. Cross-linking agents are alkylating agents and they react with nucleophilic centres in the DNA. The most nucleophilic parts of the DNA bases are the ring nitrogens. A number of chemicals induce cross-link formation and these include Mitomycin C, nitrogen mustard, sulphur mustard and photoactivated psoralens. Cross-links can also form as minor DNA damage products of UV radiation (~λ254nm) and ionizing radiation.

For repair of an interstrand cross-link the DNA adduct(s) must be removed from both of the DNA strands. In the *E. coli* incision-recombination model, the UvrABC endonuclease complex of the nucleotide excision repair (NER) pathway makes

incisions either side of the cross-link on one of the DNA strands. An incision is made 3′ to the initial incision on the 3' side of the cross-link of the DNA strand. At this stage three incisions have been made in one of the DNA strands, one 5′ to the cross-link and two 3′ to the cross-link. The oligomer between the pair of 3′ incisions is lost and the resulting gap is the substrate for RecA-mediated recombinational repair. Invasion by the homologous DNA strand during recombinational repair displaces the cross-linked oligomer. The oligomer remains attached to the other DNA strand via the cross-link. The cross-link-attached oligomer is then removed from the other DNA strand by a second round of NER and the gap generated is repaired by normal repair synthesis prior to ligation. This repair model is conservative due to the recombinational repair (Sladek *et al.*, 1989; Van Houten, 1990). In *S. cerevisiae*, repair of interstrand cross-links involves the production of a DSB. NER proteins make incisions either side of the cross-link on both DNA strands resulting in a DSB and release of the oligomers attached by the cross-link. Repair then involves the recombination proteins RAD51 and RAD52 (Dardalhon and Averbeck, 1995; Greenberg *et al.* 2001; Wang *et al.* 2001). Removal of cross-links in higher eukaryotes is not clearly understood.

In mammals interstrand cross-link removal is dependent upon the ERCC1/XPF complex. ERCC1/XPF deficient cells are hypersensitive to cross-linking agents while cells deficient in other NER genes are less sensitive (Hoy *et al.*, 1985). Two possible mechanisms have been proposed and these differ in a number of ways. The first mechanism is an error-prone translesion synthesis mechanism in which dual incisions are made, flanking the cross-link, on one of the DNA strands. DNA polymerase would then resynthesize DNA to fill the gap and the cross-link and attached oligomer would be removed by NER. In this mechanism the role of ERCC1/XPF is as part of the NER process. The second mechanism is a conservative repair mechanism in which the ERCC1/XPF complex, independent of its role in NER, makes dual incisions on one strand that flank the cross-link. Strand transfer from the intact duplex (the sister chromatid) and ligation is followed by dual incision to release the cross-link from the other DNA strand and then repair synthesis takes place. The requirement for the sister chromatid in the second mechanism means it

could only occur during DNA synthesis, while the first mechanism could operate at any time. It is possible that both mechanisms are used as the role of ERCC1/XPF in NER does not preclude an independent role for the ERCC1/XPF complex (Barre *et al.*, 1999; Greenberg *et al.*, 2001).

ERCC1 and XPF deficient mammalian cells are hypersensitive to cross-linking agents so these proteins are definitely required for interstrand cross-link removal. However, there is evidence to support both proposed mechanisms for interstrand cross-link removal. The evidence for the error-prone mechanism in which ERCC1/XPF acts as part of the NER process to remove the cross-link is as follows. The ERCC1/XPF complex can only make an incision at a cross-link if one side of the cross-link is attached to ssDNA. Therefore, it is likely that ERCC1/XPF acts as part of NER to remove the cross-link-attached-oligomer after incisions have been made to release the oligomer by another mechanism (Kuraoka *et al.*, 2000). *ERCC1* and *XPF* mutants, as well as other NER mutants, induce DSBs at the same rate as wild type when measured by PFGE assay, indicating that the ERCC1/XPF complex is not required for DSB formation during cross-link repair (De Silva *et al.*, 2000).

There is also evidence to support the second, conservative repair mechanism in which ERCC1/XPF acts independently of NER. Mammalian cells with mutations in NER genes other than *ERCC1* and *XPF* are only two to three fold more sensitive to cross-linking agents than wild type while cells deficient in *ERCC1* or *XPF* are hypersensitive, suggesting that only the *ERCC1* and *XPF* gene products are crucial for removal of interstrand crosslinks (Hoy *et al.*, 1985). The proposed repair mechanism would depend upon the presence of a sister chromatid for recombination and chromatids are only available at the G2 phase of cell replication. Cross-linking agents do result in arrest of the cell cycle at G2, supporting this hypothesis (Bridges and von Wright, 1981). Because NHEJ mutants have wild type sensitivity to the cross-linking agent mitomycin C, any double strand breaks formed during cross-link removal cannot be a substrate for NHEJ, suggesting that DSBs formed during cross-link removal are unusual in some way. An *in vitro* experiment demonstrated that the ERCC1/XPF complex could remove an interstrand cross-link. The complex could

'unhook' one side of the cross-link by making dual flanking incisions and the other side of the cross-link could then be removed by the same process (it would also act as a substrate for NER) (Kuraoka *et al.*, 2000). The cross-link 'unhooking' rate of *ERCC1* and *XFP* mutants is much slower than other NER mutants and wild type when measured by comet assay, suggesting that ERCC1 and XPF are required for this process independently of other NER proteins (de Silva *et al.*, 2000).

Two lines of evidence suggest that there may be a third, minor, mechanism that does not require ERCC1/XPF. Firstly, survival curves of *Ercc1* null fibroblasts indicated an interstrand cross-link resistant population. Secondly, mouse *Ercc1* mutants are more sensitive to the cross-linking agent mitomycin C than hamster mutants, suggesting that a pathway in hamster mutants is providing the increased resistance to interstrand cross-links in comparison to mice (Melton *et al.*, 1998; Weeda *et al.*, 1997). An ERCC1/XPF independent mechanism could involve the activity of glycosylase to 'unhook' the cross-link (Bridges and von Wright, 1981).

The 60-fold increased sensitivity of *XRCC2* and *XRCC3* (RAD51 homologues) mutants to mitomycin C suggest that these proteins are involved in the removal of cross-links (Cui *et al.*, 1999; Liu *et al.*, 1998). Because comet assays show that these mutants 'unhook' interstrand cross-links at the same rate as wild type the proteins must be involved in repair of cross-links downstream of the 'unhooking' step (de Silva *et al.*, 2000). A model for cross-link removal has been proposed that incorporates evidence of the involvement of ERCC1/XPF and XRCC2 and XRCC3. In this model the interstrand cross-link halts replication and an unknown process makes a DSB. XRCC2 and XRCC3 dependent recombination is initiated but the presence of the cross-link stalls this process. The cross-link is 'unhooked' by ERCC1/XPF (possibly resulting in a DSB if incisions are also made to free the cross-link completely) allowing recombination and repair to be completed (de Silva *et al.*, 2000). *Ercc1* knockout mice (section 1.12.8) die prematurely from liver failure and the accumulation of DNA damage in the liver of the knock-out mice, supports the role for *Ercc1* in recombinational repair of endogenous DNA damage (Chipchase *et al.*, 2003; Núñez *et al.*; Weeda *et al.*, 1997).

1.7 Base excision repair

Base excision repair (BER) is a pathway for the removal of damaged or inappropriate DNA bases (reviewed by Friedberg, Walker and Siede, 1995; Wood, 1996; Lindahl *et al.*, 1997; Fortini *et al.*, 2003). The pathway is so named because free bases are produced following excision. BER removes small DNA adducts, fragmented bases oxidised and reduced bases and apurinic/apyrimidinic (AP) sites. The damage repaired by this pathway results from a range of sources. Methylation causes formation of DNA adducts (methyl groups), oxidised, reduced or fragmented bases may arise due to the effects of free radical attack (due to normal metabolic processes) or IR. BER also repairs apurinic /apyrimidinic (AP) sites. Such sites may arise due to spontaneous hydrolysis or oxidative damage. There are two BER pathways, one for the repair of single bases (short patch BER) and one for tracts of bases (long patch BER). Although most of the proteins involved in BER are common to both pathways, some are specific to a particular pathway.

In short patch BER, recognition of the damaged base and initiation of BER is performed by one of several DNA glycosylases, with specific glycosylases targeting different types of base damage. For example, uracil is an inappropriate base to find in DNA and the glycosylase that recognises uracil is uracil-DNA glycosylase. In contrast, methylated adenine is specifically recognised by 3-methyladenine-DNA glycosylase. Following recognition of the base by the glycosylase, the N-glycosylic bond between the base and the deoxyribose-phosphate backbone is hydrolysed (Dianov and Lindahl, 1994; Frosina *et al.*, 1996; Lindahl, 1995; Memisoglu and Samson, 1988). This releases the base, forming an AP site (which is also a type of DNA damage). The next step in repair is removal of the AP site. REF1, an AP endonuclease, is required to make an incision by hydrolysis of the bond in the phosphodiester backbone 5′ to the AP site (Demple and Harrison, 1994; Sancar and Sancar, 1988; Wallace, 1988). Complex glycosylases have intrinsic AP endonuclease activity but the nick resulting from REF1 is required for repair synthesis to take place. The action of the AP endonuclease results in a strand break with a 5′ terminal deoxyribose-phosphate moiety (dRP). Before repair synthesis can take place this 5′ terminal deoxyribose-phosphate moiety must be removed. During

short-patch BER, the deoxyribose 5′-phosphatase (dRPase) activity of DNA polymerase β (pol β) removes the dRP and pol β then performs repair synthesis. Ligation in short patch BER is performed by DNA ligase III and XRCC1.

The long patch BER (2 to 10 nucleotides) pathway uses the same proteins as short patch repair until after the production of incision 5′ of the AP site. In long patch BER the dRP moiety, along with several downstream nucleotides, is removed by flap-endonuclease (FEN1). The damaged tract being repaired by long path BER is displaced and this displaced parental tract becomes a 5′ terminal overhanging region (a flap). FEN1 binds and tracks along the overhanging region from the 5′ end until it reaches the cleavage site where it makes an incision that results in release of the overhanging region. Proliferating cell nuclear antigen (PCNA) interacts with FEN1 to enhance its binding and cleavage ability. p21 is thought to then displace FEN1 from PCNA, releasing FEN1 from the replication fork so that it can perform removal of the dRP. In short patch BER the pol β performs this task but in long patch BER the replicative polymerases (pol δ/ε) are preferentially utilised and these polymerases lack the intrinsic dRPase activity of pol β. The ligation step of long patch BER is performed by DNA ligase I. Pol β may also have a role in long patch BER because FEN1 stimulates strand displacement repair synthesis by pol β but subsequent recruitment of DNA ligase I halts pol β activity, limiting repair synthesis to a single nucleotide (Fortini *et al.*, 2003; Klungland and Lindahl, 1997).

1.8 Mismatch repair

Mismatched bases are introduced into DNA by a number of processes. There are two obvious ways of introducing mismatches. The first is by incorrect base incorporation during DNA synthesis and this is likely to occur opposite bases that have been damaged by oxidation or UV (Earley and Crouse, 1998). The second is during DNA recombination if there are differences between the sequences of the DNA strands that form the heteroduplex (such differences may arise due to divergent evolution) (Wagner and Meselson, 1976). Mismatches may also occur when imperfect palindromes in the DNA sequence form a hairpin structure and when deamination of methylated cytosine (5-methylcytosine) in a G•C pair leads to

mismatching of a T opposite the G. There are two important differences between mismatch repair (MMR) and other DNA repair processes; that of *damage recognition* and *repair*. In mismatch recognition the mismatched nucleotide need not be chemically altered/damaged but is in the wrong structural location. In addition, depending upon the source of the mismatch there may be a requirement for an ability to distinguish which nucleotide in a mismatched pair is the mutation. For example, in the case of a mismatch introduced during DNA synthesis the newly synthesised strand contains the error while the parental strand must be used as a repair template. The genes involved in MMR are highly conserved between prokaryotes and eukaryotes and there is interaction between proteins of NER and MMR. The size of the excised DNA containing the mismatched nucleotide may either be <10 nucleotides (short-patch repair) or long-patch repair (as large as 10^3). The mismatch repair process is reviewed in detail in Friedberg, Walker and Siede, 1995.

1.8.1 Recognition and repair of mismatched bases in prokaryotes

When DNA synthesis or recombination produces a mismatch the genetic fidelity of the code is threatened. The incorrectly matched base of the pair must be identified and excised so that the correct sequence is used as a template for re-synthesis. If the base from the original sequence were to be incorrectly excised then repair would lead to permanent fixation of the mutation. In prokaryotes discrimination between DNA strands takes place following replication using the methylation state of the newly synthesised DNA. This can occur because of a delay between synthesis and methylation of DNA. In this way, mismatches in the daughter strand of DNA can be identified after replication because the daughter stand is under-methylated relative to the original, parental strand. Such MMR is referred to as methyl-directed MMR. A number of genes are involved in methyl-directed MMR and these are conserved between prokaryotes and eukaryotes. When a mismatch occurs the sequence surrounding the mismatch and the type of mismatch affect how quickly it is repaired. The way that mismatches are repaired means that certain types of mismatch can have more serious consequences in terms of maintenance of genetic fidelity than others. For example in a G•T mismatch, where a T had been substituted for a C, repair could lead to reversal of the mismatch (G•C>G•T>G•C) or production of an A•T transition mutation (G•C>G•T>A•T); such mismatches are repaired very efficiently because of

the danger to genetic fidelity. The second type of mismatch is one that could result in a swap of DNA code between strands, a transversion mutation. Transversions can result from repair of identical mismatched bases e.g. where the G has been incorporated in place of a C in a G•C pair, G•G may be repaired back to the original sequence or the parental G may be swapped for a C resulting in a transversion (G•C>G•G>C•G).

Recognition and repair of mismatched bases in prokaryotes involves a number of steps and proteins. Recognition, unwinding, excision, synthesis and ligation are required if a mismatch is to be repaired. In *E. coli* mismatches are recognised by the proteins MutS, MutL and MutH (Rydberg, 1978). Mismatches are initially recognised and bound by MutS (primary recognition). Due to the ability of MutL to bind MutS-bound-heteroduplex-mismatches it is thought that MutL is the secondary recognition factor in MMR. MutH has the ability to distinguish between strands on the basis of their methylation state (hemimethylated strands) and also has some endonuclease activity, meaning that it can identify the parental strand and cut the daughter strand 5′ of the mismatch. The methylation state of a specific sequence of the parental DNA, a GATC sequence where the A is methylated, allows the MutH to recognise the parental strand in comparison to the daughter strand and the GATC signal need not be close to the mismatch for recognition by MutH (a GATC-specific endonuclease). Following recognition of the mismatch and strand discrimination the DNA around the mismatch needs to be unwound and this unwinding is performed by DNA helicase II (the protein encoded by the gene *uvrD*$^+$). Single-stranded-DNA-binding protein (SSB) is also required for excision of the mismatch. Exonucleases I, VII and RecJ are required for excision of nucleotides and repair synthesis is performed by DNA polymerase III holoenzyme. It is likely that DNA ligase is required to seal the nicks following resynthesis and the co-substrate for DNA ligase, NAD^+ is known to be required for MMR to take place (reviewed in Friedberg, Walker and Siede, 1995).

1.8.2 Eukaryotic homologues of prokaryotic MMR proteins

Homologues of the prokaryotic MMR genes have been identified in eukaryotes and are highly conserved, indicating the importance of the process. In yeast (*S.*

cerevisiae) there are six homologues of the *E. coli* mutS protein (MSH1-6) and five MutL homologues (PMS1-2, MLH1-3) and they take part in recombination and/or MMR following DNA synthesis. The presence of both MutS and MutL homologues is required for MMR. There is interaction between a MutS complex and heterodimers of two MutL homologues. In yeast there are multiple MutS and MutL homologues and this is thought to be due to functional specialisation. Mismatches are initially recognised and bound by a MSH2/MSH3 or MSH2/MSH6 heterodimer. Combinations of MutS homologues form heterodimers that are involved in different tasks. During MMR MLH1 and either PMS1, MLH2 or MLH3 form a heterodimer that binds MSH2-heterodimer-bound-heteroduplex-mismatches (Kolodner and Marsischky, 1999; Marsischky *et al.*, 1996). This is homologous to the situation in *E. coli* with MutL binding MutS-bound-heteroduplex-mismatches. Excision and repair of the mismatch is likely to involve NER proteins as interaction between MSH2 and NER proteins has been reported (Bertrand *et al.*, 1998).

Human homologues of the MMR genes of yeast have also been identified. A gene called *Duc-1* in humans is thought to encode a protein very similar in amino acid sequence to the *E. coli* MutS (encoded by *mutS*$^{+}$) but has not yet been shown to have a role in MMR. There are two human MutS heterodimer complexes, hMutSα and hMutSβ. hMutSα is composed of hMSH2 and hMSH6 (homologues of MutS) and is the recognition factor for small mismatches. hMutSβ is composed of hMSH2 and hMSH3 (homologues of MutS) and is the recognition factor for larger mismatches. Once the damage has been bound by one of the hMutS complexes, the hMutL complex is recruited to the site. This heterodimer is composed of hMLH1 and hPMS2 (homologues of MutL). Mutations in the human *MLH1* and *PMS1* homologues are associated with hereditary nonpolyposis colorectal cancer (HNPCC) (Lindahl *et al.*, 1997). Following binding of hMutL, excision of the mismatch and surrounding sequence takes place and is followed by repair synthesis using the template strand. These interactions enable recruitment of other proteins involved in repair to the site.

1.8.3 MMR and recombination

As mentioned earlier, differences between the sequences of the DNA strands that form the heteroduplex during recombination can lead to mismatches. Such differences arise due to mutation or because of evolutionary divergence. Loops in the DNA form where bases are mismatched and the size of the loop structure depends upon the number of adjacent mismatches. Recombination is inhibited when mismatches are encountered during heteroduplex formation. Such mismatches inhibit recombination in eukaryotes and prokaryotes and the MMR proteins are responsible for reversal or destruction of recombination intermediates. The MutS homologues of yeast (MSH1-6) are all induced during meiosis. The MSH2/MSH6 heterodimer has anti-recombinational activity where mismatches are present. The MSH4/MSH5 heterodimer and PMS2 are required at chromosome synapsis (close pairing of homologous chromosomes) during meiosis. The MSH4/MSH5 heterodimer along with MLH1 and MLH3 are involved in crossing-over during meiosis. All these roles demonstrate the functional specialisation of the MutS and MutL homologues.

1.9 Human hereditary disorders associated with defects in NER

Study of DNA repair in mammals had been limited by the lack of available mutant cell lines but in 1968 *xeroderma pigmentosum* (XP) became the first hereditary human disorder to be associated with defective DNA damage repair. Because patients with XP are hypersensitive to UV and have a predisposition to cancer development, fibroblasts derived from the skin of a patient with this disorder were studied to discover the basis of the UV hypersensitivity. The cells were reported to be unable to perform NER following UV irradiation. This observation led to examination of the response of cells derived from patients with similar human hereditary disorders (associated with chromosomal abnormalities and increased cancer incidence) to DNA damaging agents. Cockayne's syndrome (CS) and trichothiodystrophy (TTD) were found to be associated with defective DNA damage repair. Such studies led to the identification of other genes involved in DNA damage

repair. The three diseases, XP, CS and TTD are all rare, autosomal, recessive disorders. The clinical features and genetic basis of these diseases are described below. (Friedberg, 2004; reviewed in Friedberg, Walker and Siede, 1995).

1.9.1 Xeroderma pigmentosum

Xeroderma pigmentosum is an autosomal recessive disorder. *Xeroderma* means 'parchment skin' and *pigmentosum* refers to the freckling that occurs in sun-exposed areas of the body. This is a rare disorder and incidence varies according to geographical location. The XP phenotype is severe photosensitivity and pigmentation abnormalities in exposed areas of the body, high incidence (2000-fold increase) of skin cancers with early onset, increased rate of tumour development in other UV exposed areas such as the tip of the tongue and the eyes and increased incidence of internal tumour development (20-fold increase). Neurological abnormalities are present in 20% of XP cases. The photosensitivity is due to defective NER, which results in accumulation of photoproducts that are converted to mutations during semi-conservative DNA synthesis. Neurological damage is thought to occur due to an inability to repair oxidative DNA damage in the brain. There are seven XP complementation groups (XPA - XPG) and a variant form, XPV. In XPE and variant XP NER is either normal or only slightly reduced. The genes for XPA, B, C, D, F and G have been cloned.

1.9.2 Cockayne's syndrome

Cockayne's syndrome (CS) is a rare disorder and affected individuals have a number of developmental abnormalities. The primary requirements for diagnostic features of the disease are dwarfism due to arrested growth and development (despite normal growth hormone levels) accompanied by neurological abnormalities. Other classical features of the CS phenotype are photosensitivity, cataracts, deafness, dental caries, intracranial calcifications and abnormally large facial features and limbs in proportion to body size. Five complementation groups are associated with CS but only those not also associated with XP are formally recognised as CS. The two CS complementation groups, CS-A and CS-B have the classical CS symptoms but lower photosensitivity and no predisposition to skin cancer, unlike CS associated with XP.

The genes mutated in CS-A and CS-B have been cloned. CS that is associated with XP is referred to as XP-B/CS, XP-D/CS and XP-G/CS depending on the mutations involved.

The photosensitivity of CS patients suggested that there may be an abnormality in the ability of these individuals to repair UV induced DNA damage. Fibroblasts isolated from CS patients are hypersensitive to UV radiation of λ254nm and this is due to a defect in NER. There is a defect in removal of UV induced cyclobutane pyrimidine dimers (but not other photoproducts) from the template strand in actively transcribed DNA (defective transcription coupled repair). In normal mammalian cells, RNA synthesis is depressed following UV irradiation and levels recover following repair of DNA damage. The defect in transcription coupled repair in CS cells means that the damage in the actively transcribed genes is not repaired so RNA levels cannot return to normal levels following DNA damage introduction. The defective recovery of RNA synthesis following UV irradiation is now used as a clinical method of CS diagnosis.

The defect in CS-B cells is complemented by *ERCC6*, suggesting that *CSB* and *ERCC6* are the same genes. There is a helicase motif in the predicted polypeptide of the putative *CSB* gene that has homology to transcription coupling-repair factor and this protein is thought to recruit NER proteins to transcription sites where transcription has stalled due to DNA damage in the template strand. *CSA* would likely also encode such a protein. When an individual has CS associated with XP their risk of developing skin cancer is increased when compared to cases of simple CS. This is thought to be due to the defect caused by mutation of the XP gene and may be because of an inability to repair photoproducts other than cyclobutane pyrimidine dimers.

1.9.3 Trichothiodystrophy

Trichothiodystrophy (TTD) is an autosomal recessive disorder characterised by sulphur-deficient, brittle hair (*trichoschisis* or *trichorrhexis nodosa*); fish-like scales on the skin (*ichthyosis*); mental and physical retardation and distinctive facial features (large ears and receding chin). A number of brittle-hair syndromes are associated with TTD and have acronyms which indicate the phenotype; BIDS

patients have **b**rittle hair, **i**ntellectual impairment, **d**ecreased fertility and **s**hort stature. The additional symptoms of **i**chthyosis and **p**hotosensitivity are indicated in the acronyms IBIDS and PIBIDS.

Photosensitivity is a feature of 50% of TTD cases. Although patients with photosensitive forms of TTD do not have a higher risk of developing skin cancer than normal, cells from patients with photosensitivity have defective NER with an inability to repair UV-induced DNA damage. The repair defect is associated with XP complementation group D. TTD associated with XP complementation group D, this defect is complemented by XPD. XPD has two roles; it is involved in NER and RNA polymerase II-dependent basal transcription. TTD is probably caused by defective transcription because of the association of XPD with the disorder.

There is a hypothesis that TTD is a 'transcription syndrome' and is caused by mutation of genes involved in RNA polymerase II basal transcription. There are three complementation groups associated with photosensitive TTD; TTD group A is not complemented by any of the XP genes and is likely to be due to mutation in a gene that encodes a subunit of TFIIH complex; the second group is complemented by *XPB*; the third group is complemented by *XPD*. *XPB* and *XPD* both encode subunits of the RNA polymerase II basal transcription complex TFIIH which is required for both transcription and DNA damage repair.

1.9.4 Genetic complementation groups

NER genes were identified by work on human cell lines from NER deficient patients. Similar work on NER deficient rodent cell lines had led to identification of 11 rodent complementation groups (10 from Chinese hamster and one from a murine cell line). Transfection of these cell lines with human genomic DNA led to identification and cloning of genes involved in NER. The rodent cell lines were used preferentially over human cell lines because the transfected DNA was more stably integrated into the rodent cells. As the human genes were 'cross complementing' the NER deficiency of the rodent cell lines they were given the prefix *ERCC* (excision repair cross complementing). Of the six human genes discovered by this method, *ERCC1-*

ERCC6, all except *ERCC1* were found to be mutated in human NER disorders. Several were also found to be identical to those that had already been discovered in XP, CS and TTD. Where this was the case the name of the *ERCC* gene was changed to the name associated with the human disorder. The *ERCC2*, *ERCC3*, *ERCC4*, *ERCC5* and *ERCC6* genes are referred to as *XPD*, *XPB*, *XPF*, *XPG* and *CSB*. Because *ERCC1* has not been found to be mutated in any of the patients with CS, XP and TTD it is likely that mutations in *ERCC1* are either associated with another human disorder or are developmentally lethal. The genes mentioned and the function of their encoded proteins are summarised in table 1.2.

Table 1.2 NER genes, protein function and associated human diseases.

Table shows human NER disorders, associated genes and protein function.

Gene	Corresponding ERCC gene	Associated disorder	Protein function
XPA	-	XP	DNA damage recognition
XPB	*ERCC3*	XP/XP-CS/TTD	DNA helicase, subunit of TFIIH
XPC	-	XP	Global genome repair
XPD	*ERCC2*	XP/XP-CS/TTD	DNA helicase, subunit of TFIIH
XPE	-	XP	Role in NER Gene not cloned
XPF	*ERCC4*	XP	5′ Endonuclease
XPG	*ERCC5*	XP/XP-CS	3′ Endonuclease
XPV	-	XP	Lesion bypass DNA Polymerase η
CSA	-	CS	Role in Transcription-coupled repair (Gene not cloned)
CSB	*ERCC6*	CS	Transcription-coupled repair
TTDA	-	TTD	TFIIH subunit
-	*ERCC1*	-	5′ Endonuclease

1.10 Nucleotide excision repair pathway

Because the work detailed in this thesis is based upon the role of ERCC1 in the nucleotide excision repair (NER) pathway, it is necessary to explain this pathway in greater detail than the other repair pathways mentioned. The NER pathway was discovered by study of UV sensitive mutant cells and a number of human disorders. Of all the repair pathways mentioned, NER is capable of recognising and repairing the greatest range of DNA damage. The types of DNA damage recognised by NER are UV-induced damage (CPDs and 6-4PPs), intrastrand cross-links, damaged bases and undamaged bases in the wrong structural context (mismatched bases and short loops). NER is the only repair pathway capable of removing large chemical adducts such as psoralen-thymine adducts. The lesions recognised by the NER pathway tend to be large and DNA helix-distorting. The importance of this pathway is reflected by the high degree of structural homology between NER genes between lower eukaryotes and higher eukaryotes, indicating strong selection pressure during evolution. The lack of structural homology between genes in the NER pathway of prokaryotes in comparison to eukaryotes suggests that convergent evolution has led to functional similarity. Because of the greater size and complexity of the eukaryotic genome, the NER pathway in eukaryotes involves between 20 and 30 proteins while that of prokaryotes requires only 3. However, the general mechanism of action in NER in both prokaryotes and eukaryotes is the same; *damage recognition, dual incision*, *excision*, *repair synthesis* and *ligation* (Friedberg, 1995; Sancar 1996; Wood, 1996; Lindahl *et al.*, 1997.

1.10.1 NER in prokaryotes

Early work on removal of CPDs following UV irradiation in *E. coli* revealed the presence of two separate repair mechanisms. As mentioned in section 1.5.1, the photoreactivation mechanism removes such dimers and, being dependent upon the presence of visible light, this was referred to as 'light repair'. The second mechanism capable of repairing these lesions was classified as 'dark repair' and this was the NER reaction (Friedberg *et al.*, 1995). The proteins required for NER in

prokaryotes are UvrA, UvrB, UvrC, UvrD (also known as DNA helicase II) and Pol I. UvrA and UvrB form a complex, $(UvrA)_2(UvrB)_1$, that recognises and binds DNA damage. This complex has ATP-dependent DNA helicase activity. The UvrA subunits dissociate following DNA binding, leaving the UvrB subunit bound to the DNA. UvrC is recruited by the UvrB-bound DNA and induces a conformational change that enables UvrB to make a 3′ incision 4 nucleotides from the DNA lesion. UvrC then makes a 5′ incision 5 nucleotides from the lesion. The UvrA, B and C proteins are collectively referred to as an UvrABC endonuclease although the role of UvrA is that of damage recognition rather than incision. The oligonucleotide flanked by the dual incisions is then released (excision) by UvrD. Repair synthesis to fill the gap is performed by Pol I and UvrB is released during this process and the gap is sealed by DNA ligase (Friedberg *et al.*, 1995).

1.10.2 NER in eukaryotes

Because of the homology between the genes and their roles in NER in yeast and higher eukaryotes this description of NER in eukaryotes will use the human gene nomenclature. The mammalian NER genes were largely characterised using UV sensitive rodent cell lines defective in excision repair (the excision repair cross-complementing genes) and cell lines from the seven XP complementation groups. A number of the rodent genes were found to correspond to genes mutated in XP, CS and TTD. Where this is the case, the human gene nomenclature is adopted. As mentioned earlier, there is a high degree of homology between the higher and lower eukaryotic NER genes. Table 1.3 details the relationship between homologous human, murine and yeast genes. There are two NER pathways, Global genome repair (GGR) and Transcription coupled repair (TCR). GGR removes lesions throughout the genome independent of transcription while TCR removed lesions from actively transcribed genes (Bohr *et al.*, 1985; Friedberg, 1996; Tornaletti and Hanawalt, 1999). Probably because of the imminent requirement for stalled transcription to resume, TCR removes lesions from DNA at greater rate than GGR (Mellon *et al.*, 1987). Because the transcription machinery is stalled by the presence of lesion, in TCR there is not such a requirement for the role of the proteins with

DNA damage affinity as there is in GGR (Tornaletti and Hanawalt, 1999). Probably for this reason, XPC is not essential to TCR (Venema *et al.*, 1991). For this reason, the 'bipartite recognition model' in which the genome is scanned for DNA distorting lesions and then the lesions are localised to a particular strand refers more to the case of GGR than TCR.

Table 1.3 NER genes in eukaryotes

Table of cloned human NER genes and their names in the mouse and *S. cerevisiae*

Human	**Mouse**	**Yeast (*S. cerevisiae*)**
ERCC1	*Ercc1*	*RAD10*
XPA	*XPA*	*RAD14*
XPB	*Ercc3*	*RAD25*
XPC	*XPC*	*RAD4*
XPD	*Ercc2*	*RAD3*
XPE	*DDB1-DDB2*	
XPF	*Ercc4*	*RAD1*
XPG	*Ercc5*	*RAD2*
CSA	*Ercc8*	*RAD28*
CSB	*Ercc6*	*RAD26*

As in prokaryotic NER, the stages involved in lesion excision are *damage recognition*, *dual incision*, *excision*, *repair synthesis* and *ligation*. The current model involves 25 proteins and has been reconstituted *in vitro* (Aboussekhra *et al.*, 1995). The damage recognition and dual incision stages of NER were reconstituted with RPA, XPA, TFIIH (containing XPB and XPD), the ERCC1/XPF complex and a factor designated IF7. To achieve the final stages of NER, replication factor (RFC),

proliferating cell nuclear antigen (PCNA), DNA polymerase ε and δ and DNA ligase I are required.

The type of DNA lesion recognised by the NER pathway generally causes distortion of the DNA helix. Most investigations of NER use UV-damaged DNA as the substrate. The primary lesions resulting from UV irradiation, CPDs and (6-4)PPs both cause distortion of the DNA helix, with the degree of conformational change depending upon the type and orientation of the lesion. For illustrative purposes, cyclobutane pyrimidine dimer causes unwinding of the DNA helix by 19.7° and results in a 27°-30° kink that protrudes into the major groove of the DNA. The differences between the distortions caused by the two types of UV induced intrastrand crosslinks, CPDs and (6-4)PPs, may be the reason why NER removes them at different rates, with CPDs being removed more slowly than (6-4)PPs (Gunz *et al.*, 1996; Husain, *et al.*, 1988; Pearlman and Holbrook, 1985; Ramiro *et al.*, 2004; Reardon and Sancar, 2003). In NER the DNA lesion must be excised without causing damage to the opposite DNA strand, as the undamaged strand is required as the template for repair synthesis. The DNA strand containing the lesion must, therefore, be identified following detection of the distorted DNA at the damage site. These requirements have led to the bipartite substrate discrimination model in which the general location of the lesion is first detected by the degree of disparity between the shape of the normal DNA duplex and distortion due to the lesion. Once the general location of the damage has been identified, proof-reading identifies the strand containing the lesion, initiating incision. The most widely accepted model for NER is that XPC-hHR23B, TFIIH, XPA, RPA, XPG and ERCC1/XPF form the damage recognition/dual incision complex. Following dual incision the repair synthesis/ligation complex components are recruited: DNA polymerase ε and δ, RFC, PCNA and DNA ligase I.

The order of assembly of the repair complex is reviewed in Dip *et al.*, 2004 and their interactions are reviewed in Araujo and Wood, 1999. The proteins involved in NER sequentially assemble into two complexes to perform the various stages of repair. The proteins involved have been identified and their roles are broadly understood.

The order of assembly, especially the order that the damage recognition proteins assemble, is not yet clear. It may be that the initial order of assembly varies depending upon the substrate as several of the proteins have been implicated in the initial damage recognition step of NER. XPA and RPA form a recognition complex. There is some specificity in the type of damage these proteins have affinity for and this has led to three models; "XPC first" (Evans *et al.*, 1997; Naegli, 1995; Sugasawa *et al.*, 1998; Volker, 2001), "RPA/XPA first" (Wakasugi and Sancar, 1998) and "RPA first" (de Laat *et al.*, 1998). In these models, which refer to the GGR pathway, either XPC (bound to hHR23B - the human homologue of RAD23), RPA/XPA or RPA alone binds to the damage and subsequently recruits the other components of the NER complexes. There is evidence to support all three models so it is likely that depending upon the type of damage and whether it is encountered during GGR or TCR, either XPC or RPA/XPA or RPA is the primary molecular recognition component. Although XPC has a higher affinity for most types of DNA damage than XPA and RPA no affinity for CPDs has been detected (Batty *et al.*, 2000; Reardon *et al.*, 1996; Sugasawa *et al.*, 2001). RPA, but not XPA nor XPC, is recruited to psoralen adducts. It has been suggested that all three of the proteins function cooperatively to identify damage and recruit factors for repair to the site (Reardon and Sancar, 2003). The role of XPE in NER is unclear. Damaged DNA binding protein (DDB) has an affinity for damaged DNA and the heterodimer has a 48kDa subunit encoded by the *XPE* gene. This subunit is mutated in all known cases of XP complementation group E. DDB has a high affinity for CPDs and when it binds to these lesions it induces a 55° bend in the DNA helix (Fujiwara *et al.*, 1999; Hwang and Chu, 1993; Payne and Chu, 1994; Reardon *et al.*, 1993). It may be that this results in binding by XPC (or XPA) (Dip *et al.*, 2004).

Following detection of damage, the basal transcription factor TFIIH is recruited. This factor has nine subunits, two of which are the DNA helicases XPB and XPD (van Vuuren *et al.*, 1994). TFIIH probes and unwinds the helix around the lesion, RPA binds the undamaged strand and the localised unwinding of the DNA allows access to the lesion by the endonucleases. Aside from the ability of TFIIH to unwind the DNA at the lesion, it is possible that this factor also has some DNA damage

recognition capability; the XPD subunit TFIIH is located very close to the lesion in the fully assembled recognition/excision complex (Reardon and Sancar, 2002; Ramiro *et al.*, 2004). TFIIH interacts with XPA and recruits the endonuclease, XPG. The endonuclease, XPG interacts with RPA and XPA recruits the endonuclease heterodimer ERCC1/XPF. XPG makes the incision 3′ of the lesion while ERCC1/XPF makes the 5′ incision. XPG makes an incision 6-9 nucleotides 3′ to the lesion (O'Donovan *et al.*; 1994, Huang *et al.*, 1992; Moggs *et al.*, 1997). ERCC1/XPF then makes an incision 16 to 25 nucleotides 5′ to the lesion (Sijbers *et al.*, 1996). These incisions are not, therefore, equidistant from the lesion. Following dual incision the 22-30 base oligonucleotide is excised, leaving a single stranded gap bound by RPA (Mu *et al.*, 1996). RPA is then involved in dissociation of the damage recognition and excision proteins and recruitment of the repair synthesis proteins (Reidl *et al.*, 2003). RFC binds to the RPA-bound single stranded gap and recruits PCNA, a sliding clamp for DNA polymerase ε and δ (Shivji *et al.*, 1995). Repair synthesis of the gap follows and the newly synthesised DNA is ligated by DNA ligase I.

As mentioned earlier, TCR does not require XPC because the DNA damage is located when the RNA polymerase II holoenzyme is stalled by the DNA lesion. Detection of a lesion in this way leads to recruitment of XPA, RPA, CSA and CSB. Some mismatch repair proteins and the UV^{S}S protein (*UV^{S}S* is mutated in a UV-sensitive syndrome distinct from XP and CS) are also required for TCR (Le Page *et al.*, 2000; Mellon *et al.*, 1996; Spivak *et al.*, 2002). TFIIH is able to move from its location in the RNA polymerase II enzyme to take part in recruitment of the endonucleases and NER proceeds as in GGR from the dual incision stage onwards (Dip *et al.*, 2004; Drapkin and Reinberg, 1994).

1.11 Cloning and characterisation of ERCC1

Because *ERCC1* is the focus of this thesis, the gene and its protein will now be described in detail. As described earlier, the ERCC genes (*ERCC1-6*) were identified in genetic complementation studies of 11 UV-hypersensitive rodent cell lines. Subsequent analysis revealed that all but one of the ERCC genes

complemented the human XP and CS complementation groups. *ERCC3,4* and *5* complemented XP groups and *ERCC 6* complemented CS group B. The fact that humans exist with these disorders indicates that deficiency in ERCC genes 2-6 is not lethal. No human disorder has yet been associated with deficiency in *ERCC1*, indicating that *ERCC1* deficiency is embryonic lethal in humans or deficiency is compatible with survival but the associated human disorder has not yet been identified (van Duin *et al.*, 1989). Since the NER diseases are very rare, this is a possibility. However, because XP patients completely deficient in NER survive it may be that an accessory role of *ERCC1* is so important to viability that deletion of *ERCC1* is incompatible with survival in humans. *Ercc1* knockout mice perish shortly after birth due to liver failure that has been attributed to the lack of Ercc1 leading to a deficiency in recombinational repair.

ERCC1 complements the mutation in CHO43-3B cells (Westerveld, 1984). These cells are UV hypersensitive and the *ERCC1* mutation is a missense mutation in the 98th reside (Hayashi, 1998). In humans the gene has 10 exons and spans 15kb of chromosome 19. The mouse gene was cloned using human cDNA and this gene also has 10 exons, the gene is located on chromosome 7 (van Duin *et al.*, 1988). There is 82% sequence homology between the mouse and human gene and both are driven by a promoter that lacks classical promoter elements (discussed in introduction to chapter 3). The 3′ end of the gene overlaps with the 3′ end of the antisense *ERCC1* gene, *ASE1*, the function of which is unknown (van Duin *et al.*, 1989). The gene is expressed constitutively at low levels throughout the body with the primary transcript being 1.1kb (with the exception of a novel skin transcript of 1.5kb, described later). Alternative splicing and polyadenylation results in production of additional transcripts including a 1kb transcript lacking exon 8 that does not encode a repair proficient protein (van Duin *et al.*, 1986), a 4.3kb and a 3.8 kb transcript. The constitutive expression of this gene is in common with the other NER genes and the transcript was detected at all developmental stages studied (van Oostrom *et al.*, 1994), although recent work in this laboratory demonstrated that the transcript is absent in early mouse embryos or is expressed at levels too low to detect by northern blot and RT-PCR (unpublished data). Studies to date indicated that the gene was not

induced by UV. The gene is highly conserved in evolution; there is even a homologue in plants that can complement mitomycin C sensitivity of *ERCC1* deficient mammalian cells (Xu *et al.*, 1992).

1.11.1 The ERCC1 protein

The ERCC1 protein in humans is 297 amino acids long while in mice it is 298 and the weight is 33kDa. There is 85% homology between the mouse and human proteins overall. The first 100 amino acids of the C-terminus have 70% homology, the next 100 have 97% homology and the final 98 amino acids have 89% homology. The middle region of the protein is, therefore, the most conserved region and a putative DNA binding domain has been assigned to amino acids 134 to 156 in this area (Hoeijmakers *et al.*, 1986). The middle part of the *S. cerevisiae* RAD10 protein shares 35% homology with 110 amino acids of the C-terminal of the mammalian *ERCC1* protein and in the middle of this homologous region there is 56% homology over a stretch of 25 amino acids. In RAD10, the homologous region encodes a helix-loop-helix DNA binding motif. There is no homology in the N-terminal region (van Duin *et al.*, 1986). RAD10 and RAD1 associate in yeast to perform the 5′ incision in NER, the role performed by ERCC1/XPF in mammalian NER. There is also a homologue of *ERCC1* in *S. pombe*, SWI10 and this associates with the *S. pombe* homologue of XPF, spRAD16. The RAD10 and RAD1 has a role in recombination in yeast (Tomkinson *et al.*, 1993, Bardwell *et al.*, 1994); *ERCC1*/XPF seems to be involved in recombination in mammals (Chipchase *et al.*, 2003). Deletion of 5 residues from the C-terminus of the mammalian ERCC1 results in complete inactivation of the protein (Sijbers *et al.*, 1996). A putative nuclear localisation signal has been assigned to amino acids 17 to 23 of the N-terminus and the mammalian *ERCC1* protein has 83 N-terminal amino acids (not found in yeast) which has 38% homology to the C-terminus of the *E. coli* NER protein UvrC (van Duin *et al.*, 1988; Hoeijmakers *et al.*, 1986). However, deletion of 91 N-terminal amino acids of ERCC1 has no effect upon the activity of the protein.

1.11.2 Interaction between ERCC1 and XPF.

The first indication that ERCC1 and XPF act as a complex in NER came from the inability of cell-free extracts from *ERCC1* and *ERCC4* deficient rodent cell lines (and human XPF cells) to complement repair deficiency *in vitro*. At this stage, *XPF* and *ERCC4* were suspected to be equivalent and this was later proven (Sijbers *et al.*, 1996). The free ERCC1 protein is unstable (Yagi *et al.*, 1997); in cell-free extracts from XPF deficient cells there is no available ERCC1 to bind to XPF and complement the ERCC1 deficient cells, since the lack of XPF in the XPF deficient cells means the ERCC1 is unstable. The association between the proteins was again suggested by an experiment in which ERCC1 antibodies were added to cell-free extracts from wild type (wt) cells to deplete the ERCC1 levels. This could not complement ERCC4 or XPF deficient cells, indicating that removal of ERCC1 resulted in removal of XPF (ERCC4) (van Vauren, 1993). The correcting activities for ERCC1, ERCC4 and XPF co-purified as a high molecular weight species ~100-120kDa and this complex was able to co-correct the three complementation groups *in vitro*). The complex was found to contain the ERCC1 protein (33kD) and a 112kDa protein (ERCC4 or XPF) with endonuclease activity (Biggerstaff *et al.*, 1993). ERCC1 and XPF (ERCC4), therefore, formed a heterodimer capable of endonuclease activity in NER. The domains responsible for stable binding in the complex were located using C- and N-terminal truncated recombinant ERCC1 and XPF proteins. The binding domains of the ERCC1 and XPF proteins were localised to residues 224- 297 and 814-905 respectively (de Laat *et al.*, 1998). The XPF binding region of ERCC1 is in the region of homology to the *E. coli* UvrC protein mentioned earlier, so the interaction domain between ERCC1 and XPF does not correspond to RAD1/RAD10 in *S. cerevisiae* as this region is not present in the yeast ERCC1 yeast homologue. There is also an interaction domain that enables ERCC1 to bind to XPA, one of the damage recognition proteins of NER and therefore required for recruitment and positioning at the damage site. The ERCC1/XPF complex was shown to bind to an XPA affinity column (Park and Sancar, 1994) and truncated recombinant proteins (as with the XPF and ERCC1 experiment) localised the ERCC1 and XPA interaction domains to residues 92-119 and 75-114 respectively (Li *et al.*, 1994). This explains the NER deficiency of the Chinese hamster cell line (43.3B),

complemented by ERCC1, as the mutation in this cell line (the 98th residue) lies within the XPA interaction domain; the mutation prevents interaction between ERCC1 and XPA during NER.

1.12 Animal models of NER disorders

A number of mice with inactivating mutations in DNA repair pathway genes have been produced with the intention of gaining insight into the DNA repair mechanisms. The rarity of human DNA repair disorders such as XP, CS and TTD mean that *in vivo* mouse models are of great importance to furthering understanding of the role of DNA repair genes. Mice with deficient DNA repair pathways (such as NER and MMR) are useful for the study of carcinogenesis and mutagenesis, as well as therapy development, in terms of the disorder they are developed to model. They are also useful for general studies of potentially carcinogenic and mutagenic substances (Wijnhoven and van Steeg, 2003). A database of mouse strains carrying targeting mutations in genes affecting biological responses to DNA damage initiated in 1997 is now at version 5 and this indicates the high rate of mouse knockout production and the interest in this approach for modelling human DNA repair disorders (Friedberg and Meira, 2003). Mice produced with mutations in DNA repair genes often exhibit a number of phenotypic abnormalities found in the associated human disease, serving as good models. However, this is not always the case as the consequences of the gene inactivation in the mouse often have unexpected results. These differences can often be attributed to differences between the physiology/developmental processes of the mouse and human due to divergent evolution. For example, deficiency in the enzyme hypoxanthine phosphoribosyltransferase (HPRT) in humans results in a disease called Lesch-Nyhan syndrome, a severe neurological disorder resulting in mental retardation and self mutilation (Lesch and Nyhan, 1964). In contrast, mice lacking in the gene are phenotypically normal (Hooper *et al.*, 1987) because mice are more reliant upon adenine phosphoribosyltransferase (APRT) than HPRT in their purine salvage pathway than humans (Wu and Melton, 1993). Mouse models are a useful tool for study of NER disorders but care is required when results of gene inactivation in the mouse are extrapolated to humans. The *Ercc1* null mouse was the

first *in vivo* model for defective NER. Since this mouse was developed a number of other NER mutant mice have been developed.

1.12.1 *Csa* null mice

Csa null mice are viable (van der Horst *et al.*, 2002). The null mice have phenotypic features in common with human CS patients. The mice are UV hypersensitive as demonstrated by the UV hypersensitivity of embryonic fibroblast cell lines derived from the mice. The mice have defective TCR but normal GGR. The mice also exhibit pronounced age-dependent loss of retinal photoreceptor cells. However, the mice do not develop the severe developmental and neurological abnormalities associated with human CS patients. In addition, the null mice develop skin tumours and this is not a feature of CS in humans. The increased risk of skin cancer development in the mice may reflect less efficient global genome repair than in humans in the absence of TCR.

1.12.2 *Csb* null mice

Csb null mice are viable (van der Horst *et al.*, 1997). Mice were produced with a mutation similar to one discovered in a human CS patient; a premature stop codon was inserted into the fifth exon of the *Csb* gene. The mice do produce a mRNA fusion transcript but this is not translated into protein. The null mice share a number of phenotypic features with CSB patients; the mice had reduced growth compared to wild type animals and behavioural analysis suggested the mice have a mild neurological abnormality. The mice were UV hypersensitive and fibroblast cell lines derived from the mice had defective TCR but GGR levels were normal. All these symptoms are features of mild CS in humans and in this respect the mice are good models for the disorder. However, the mice had some features that are not found in human CSB patients; the mice have a predisposition to developing skin cancer, do not exhibit impaired sexual development and do not exhibit neurological dysmyelination.

1.12.3 Xpa null mouse

Human XPA patients present the most severe phenotype of the seven XP complementation groups. Generation of *Xpa* null mice has been reported twice (de Vries *et al.*, 1995; Nakane *et al.*, 1995). Two different methods were used to generate the mice; in one group's mouse a *neo* cassette was inserted into exon 4 and the other group deleted exons 3 and 4. Both mice were shown to be complete 'knockouts' by northern or western blot analysis. The mice were viable and fertile, with a normal appearance up to 13 and 18 months and did not exhibit the progressive neurodegeneration associated with XPA group patients. The mice were NER deficient and this was demonstrated by the hypersensitivity of null mouse-derived fibroblast cell lines to UV and genotoxic compounds. The residual repair of DNA damage was extremely low (<5% in the mouse) and this is comparable to the level of repair detected in XPA patients (<2%). Both groups reported that their mice had a markedly increased rate of UVB induced skin and eye cancer when compared to wild type littermates. 21% of mice aged over 1 year were later reported to have increased rates of spontaneous liver tumour development (de Vries et al., 1997). The *Xpa* null mice are a good model for study of skin, eye and liver cancer development observed in human XPA patients but cannot be used as a model for research of the neurodegeneration observed in 20% of human patients.

1.12.4 *Xpc* null mice

Xpc null mice are viable and have no phenotypic abnormalities. However, cell lines derived from the mice are hypersensitive to UV. Mice have an increased risk of developing skin cancer as well as UV-induced pathological skin and eye changes. These symptoms are consistent with XPC patients (Sands *et al.*, 1995).

1.12.5 *Xpd* mutant mice

Complete deletion of *Xpd* results in a pre-implantation embryonic lethal phenotype (de Boer *et al.*, 1998). This is thought to be due to the inability of cells lacking XPD to perform basal transcription due to inactivation of the TFIIH transcription factor. Because complete 'knockouts' were not viable, mice containing a point mutation in

the *Xpd* gene were generated. The location of the point mutation was chosen to simulate a mutation discovered in a patient with the photosensitive form of TTD. The mutation was introduced via 'gene-cDNA fusion targeting' in which human cDNA containing the mutation replaced the homologous sequence at the 3' end of the mouse gene. Mice generated using this modified gene share a number of phenotypic features with human TTD patients; the mice have brittle hair, reduced growth, skin abnormalities, reduced fertility and reduced lifespan. Human TTD patients generally develop neurological abnormalities over time and the occasional tremors reported in the mice suggest that they may also have neurological abnormalities. The increased hypersensitivity to UVB measured in fibroblasts derived from the mice (1.5 fold higher than wt) matched the level measured in human fibroblasts derived from TTD patients. Although the mice are good models for most of the symptoms of TTD, the predisposition towards UVB and chemically induced skin cancer development observed in the mice does not conform to the human TTD phenotype. As in the case of CSB mice, it was suggested that this could be due to less efficient GGR in mice than humans.

1.12.6 *Xpf* mutant mice

Because the proteins encoded by *XPF* and *ERCC1* form a heterodimer it was assumed that *Xpf* null mice would exhibit the severe phenotype of *Ercc1* mutants. A group attempting to create a mouse with only partial inactivation of the *Xpf* gene, for a long term study of UV carcinogenesis and to investigate the role of *ERCC1*/XPF in immunoglobulin class switch recombination (CSR), accidentally generated a complete *Xpf* knockout (Tian, 2004). A mutation was introduced into the gene that resulted in the formation of a premature stop codon. However, this mutation resulted in complete inhibition of *Xpf* expression. The mice had exactly the same phenotype as *Ercc1* deficient mice, the severe postnatal growth defect, abnormalities in the nuclear size of liver cells and early death at 3 weeks after birth. Embryonic fibroblasts cell lines derived from the mice were hypersensitive to UV and mitomycin C (cross-linking agent). Human XPF patients show mild symptoms of photosensitivity and occasional neurological abnormalities but do have low levels of XPF protein.

1.12.7 *Xpg* mutant mice

Xpg null mice were produced as a model for human XPG/CS. The mice were produced by insertion of a *neo* cassette into exon 3 of the *Xpg* gene. The mice had growth retardation that became increasing severe with time, premature death by day 23 after birth, progressive emaciation and weakness, developmental retardation of the brain leading to progressive ataxia and decreased activity levels. The mice share a number of phenotypic features with human CS and XPG patients (the small brain size, developmental retardation and progressive neurological dysfunction). Because XPG and early onset CS patients are usually have truncated XPG protein rather than complete lack of protein a second type of mutant mouse was generated by the same group to more closely mimic this situation. These mice had a truncation of the protein C-terminus due to genomic *Xpg* DNA being replaced with a mutated cDNA fragment. These mice had growth retardation and short life span, matching the symptoms of human XPG/CS patients (Sun *et al.*, 2003; Shiomi *et al.*, 2004

1.12.8 *Ercc1* null mice

Because no human disorder is associated with mutation of the *ERCC1* gene it had been suggested that *ERCC1*-deficiency resulted in an embryonic lethal phenotype. *Ercc1*-deficient mice were successfully developed in the Melton laboratory and these were the first mouse *in vivo* model for defective NER (McWhir *et al.*, 1993). The *Ercc1* gene was inactivated by insertion of a *neo* cassette into exon 5 and the mice were produced using the embryonic stem cell method. The *Ercc1* null pups have a range of severe phenotypic abnormalities including severe runting (weighing 20% less than wild type littermates) and death within three weeks of birth (prior to weaning) due to liver failure. The liver cells of the mice had nuclear abnormalities. These abnormalities grew progressively worse with time and the appearance of polyploid liver cells (accompanied by elevated p21 levels) was highly unusual as such high levels of polyploidy are normally only found in 1- to 2-year-old wild-type mouse livers. The slightly elevated p53 levels in the liver, brain and kidneys of the *Ercc1* null mice indicated that DNA damage was accumulating in these tissues rather than being repaired, yet increased apoptosis was not detected in these tissues. This

appearance of polyploidy in response to accumulating DNA damage may reflect an additional non-NER-related function for the Ercc1 protein and it is interesting that the elevated p21 associated with this premature polyploidy development is not a feature of polyploidy in wild type adult livers (Chipchase *et al.,* 2003, Nunez *et al.,* 2000). The general appearance of the mice mimicked that of premature ageing. The NER deficient status of the *Ercc1* null mice was demonstrated by the UV hypersensitivity of embryonic fibroblast cell lines derived from the mice. The cells were also mildly sensitive to the cross-linking agent mitomycin C and this result indicated that recombination mediated repair of interstrand cross-links is either not entirely dependent upon *Ercc1* in the mouse or that this function was not affected by the targeting. The ability of the cells to perform homologous recombination was normal. Northern blot analysis of the null mouse tissues detected an *Ercc1*/neo fusion transcript that had the potential to encode more than the N-terminus of the protein. The targeting did inactivate the NER capability of *Ercc1* in the mice but they are not a model for the human NER disorder XP (Melton *et al.*, 1998).

Two *Ercc1* null lines were subsequently independently developed by a group in Holland (Weeda *et al.*, 1997). In one 'knockout' line the *Ercc1* gene was disrupted by insertion of a *neo* cassette into exon 7 and in the second line a seven amino-acid carboxy-terminal truncation was engineered by insertion of a premature stop codon at residue 292. The mice shared the runting and premature death phenotype of the original *Ercc1* null mouse produced in the Melton Lab., but the line with the truncation had less severe symptoms. The Dutch mice had a longer life span than the original *Ercc1* knockout, 78 days in the null mice and up to 6 months in the mice with the truncation. This may be due to the fact that these mice were of a different genetic background to the original null mice or it may be due to some low levels of *Ercc1* activity. A number of extra phenotypic abnormalities were reported in the Dutch mice; nuclear abnormalities of the kidney, kidney malfunction, absence of subcutaneous fat and early onset of ferritin deposition in the spleen (which can result from increased turnover of erythrocytes). It is likely that the enhanced life span of the Dutch mice relative to the original *Ercc1* null line gave these symptoms time to develop and subsequent examination of the original *Ercc1* null line detected some of

the features reported in the Dutch mice. Embryonic fibroblast cell lines derived from the two Dutch mutants were UV hypersensitive and mitomycin C sensitive and this confirmed the NER deficient status of the mice. As with the original *Ercc1* null mouse, the Dutch mice could also potentially give rise to a transcript encoding more than the N-terminal half of *Ercc1*.

The short lifespan of the original *Ercc1* knockout limited the type of *in vivo* analysis that could be performed. In order to investigate the effect of the *Ercc1* deficiency in tissues other than the liver a second mouse line was developed in the Melton Lab. with a liver specific *Ercc1* transgene. This line has a longer lifespan than the original knockout (up to 10 weeks) and the liver is normal, indicating that the cause of premature death in original mice was due to liver failure. The liver transgene mice developed kidney damage by 10 weeks.

1.13 Ercc1 expression in vitro and in vivo

1.13.1 *Ercc1* expression in mouse tissues and discovery of novel transcript

In mice and humans the *ERCC1* gene is expressed constitutively at low levels. This is in keeping with the expression pattern of the other NER genes. Earlier work in this laboratory by Dr. Selfridge revealed that the 1.1kb transcript was expressed in skin, brain, heart, lungs, liver, spleen, kidney, testes and muscle. There was some heterogeneity in the size of the *Ercc1* transcript detected in the brain and heart but a discrete, novel transcript of 1.5kb was discovered in the skin and this transcript was outwith the range of the heterogenous transcripts in brain and testes (see figure 1.1). The upstream transcript was also detected in human skin. The 1.5kb transcript is the major *Ercc1* transcript in the skin and no evidence of this transcript was found in the other tissues studied. The novel transcript is also the major transcript in human skin (Dr. Winter, Melton Lab., unpublished data). It is the novel transcript that forms the main focus of the work presented in this thesis. The 1.5kb transcript originates 400bp upstream of the normal mouse 1.1kb transcript and both the 1.1kb and 1.5kb transcripts encode the same protein. The skin-specific transcript initiates upstream of the normal transcript's promoter, indicating the presence of a second, skin-specific,

promoter. What makes the 1.5kb, skin-specific, transcript particularly interesting is that it is only present in mouse skin and it is the primary transcript in this tissue. ERCC1 is required for removal of UV-induced DNA lesions and the skin is the main tissue exposed to UV in the body. Our hypothesis is that the novel transcript confers an advantage to skin in helping to meet the higher biological demand for ERCC1 activity in this tissue. The elevated expression of *Ercc1* observed in the testes may be related to the high demand for recombination in this tissue. The *Ercc1* expression patterns of a number of cell lines were then characterised by Dr. Selfridge in order to facilitate investigation of the novel transcript's induction, promoter region and function. Skin-derived cells, keratinocytes and dermal fibroblasts both express the novel transcript; in keratinocytes the major transcript is the 1.5kb, skin-specific transcript. In fibroblasts the level of 1.1kb and 1.5kb transcripts is equal. Mouse embryonic stem cells and a CHO cell line express only the 1.1kb transcript. *Ercc1*-deficient derivatives of the keratinocyte, embryonic fibroblast and CHO cell lines were also available. The use of these cell lines is discussed in detail in chapter 3. Because of the higher biological demand of the skin for UV induced DNA damage repair, it may seem obvious that increased *Ercc1* transcription (and protein levels) would confer a survival advantage in the avoidance of UV-induced carcinogenesis. However, it is not clear why this requirement is not met by increased expression of the normal transcript. It is, therefore, of interest to investigate expression of the skin transcript in more detail.

1.13.2 Ercc1 protein levels in the skin relative to other tissues

The level of mRNA in different mouse tissues has been studied and Fig. 1.1. demonstrates typical expression patterns for a panel of mouse tissues. Both the 1.1kb and 1.5kb *Ercc1* transcripts encode a fully-functional Ercc1 protein, so both these transcripts contribute to the total protein level in the skin. Although no published data is available, unpublished data from the Melton Lab. demonstrated that the level of protein expressed in the skin is proportional to the total level of 1.1kb and 1.5kb transcript in the tissue. Because the 1.1kb transcript is expressed constitutively at low levels, it is the level of 1.5kb transcript that mainly contributes to the protein level in the skin; mouse strains expressing high levels of 1.5kb transcript have

corresponding high protein while mice with low levels of 1.5kb transcript have low levels of protein (see figure 7.1). The possible implications of the high 1.5kb transcript expression in the skin in relationship to Ercc1 protein level is discussed in detail in chapter 7. One reason for the paucity of published data regarding the relationship between transcript expression and corresponding protein levels in the mouse is that the original antibodies for western analysis were raised against human ERCC1 and did not recognise the mouse protein. The level of transcript in human skin is proportional to the level of protein. Without further analysis, no firm conclusions can be made regarding the relationship between *Ercc1* transcript expression and resultant protein level.

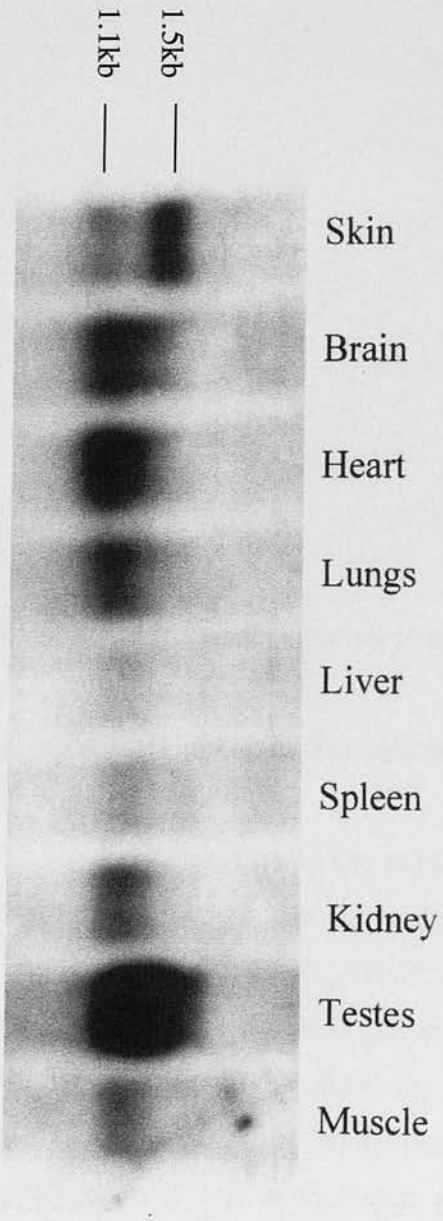
In Fig. 1.1, there is variation in the level of *Ercc1* transcript expression in the panel of mouse tissues examined. Assuming that the level of protein in these tissues is proportional to transcript level, as it appears to be in skin, the protein level is likely to be lower than skin in tissues such as the liver, spleen and muscle. The 1.1kb transcript is expressed at high levels in the heart, brain, lungs and testes; the protein level in the lungs is likely to be equal to the level in skin, while the high level of 1.1kb transcript in the brain, heart and testes is likely to lead to a higher level of Ercc1 protein in these tissue in comparison to skin. It is not known whether the heterogeneity in transcript size in the brain, heart and testes reflects expression of transcripts capable of encoding a functional Ercc1 protein. If these transcripts do encode a functional protein then the level of protein in these tissues is likely to be higher than in the skin. It is possible that the high level of transcript observed in internal tissues reflects a high biological demand for ERCC1 protein, which may relate to non-DNA damage repair roles such as recombination.

Figure 1.1 Identification of a novel *Ercc1* transcript in mouse skin

This work was performed by Dr. Selfridge, Melton Lab.

Northern blot analysis of a panel of tissue RNA from a wild type male mouse. Panel shows tissues probed with an 800bp BamHI fragment of *Ercc1* cDNA, corresponding to exons 1 to 8. The RNA loads were equal (ethidium bromide staining image could not be clearly reproduced from original).

The discrete skin-specific transcript is visible in the skin.


1.5kb
1.1kb
Skin
Brain
Heart
Lungs
Liver
Spleen
Kidney
Testes
Muscle

1.14 Project Aims

The role of the recently discovered skin-specific *Ercc1* transcript is unknown. The normal transcript is expressed constitutively at low levels in the body while the novel transcript is expressed at high levels only in the skin. Because the skin has a high biological demand for repair of UV-induced DNA damage and non-repair of UV-induced lesions can eventually lead to development of serious cancers such as malignant melanoma, it may be that evolutionary selection has resulted in high expression of the novel transcript in this tissue. The role of the skin-specific transcript required investigation. The skin-specific *Ercc1* transcript initiates upstream of the normal transcript's promoter, indicating the likely presence of a second promoter. No classical promoter elements are present in the region flanking the *Ercc1* gene but an intriguing CT repeat region had been identified. Identification of regions required for expression of the skin-specific transcript was necessary. The aims of the work presented in this thesis were;

(A) To investigate the promoter region of the skin-specific *Ercc1* transcript.

(B) To investigate the role of the skin-specific *Ercc1* transcript in the skin.

Chapter 2

Materials and Methods

2.1 Materials

2.1.1 Laboratory reagents and suppliers:

Amersham Bioscience UK Ltd.: NICK™ Column

BDH Ltd.: Grams crystal violet solution, hydrochloric acid

Biogenesis: RNA-Bee™

BIOLINE: HYPERLADDER I™

BIO-RAD: CHELEX® 100 Resin (200–400 mesh)

Boehringer Mannheim UK Ltd.: deoxyribonucleotides

Difco Laboratories: agar, bacto-tryptone, bacto-yeast extract

Fisher: boric acid, chloroform, EDTA, ethanol, glacial acetic acid, isopropanol, methanol, potassium acetate, sodium acetate, sodium chloride, sodium citrate, sodium hydroxide, trichloroacetic acid

Invitrogen™ life technologies: agarose, ultraPURE agarose electrophoresis grade, foetal bovine serum, formamide, phenol, Tris

Kimberley Clark Ltd.: paper towels

New England BioLabs Inc.: 100X purified BSA

NEN® Life Science Products: Genescreen *Plus*®

Promega: RNasin® Ribonuclease Inhibitor

Qiagen Ltd.: tip 500 columns,

Roche: Deoxynucleoside triphosphate set (dNTPs), random hexamers

Scientific Laboratory Supplies: Semperguard gloves

Sigma Chemical Co. Ltd.: β-mercaptoethanol, bromophenol blue, dextran sulphate, DTT, epidermal growth factor (EGF), ethidium bromide, ficoll, formaldehyde, glycerol, herring sperm DNA, mineral oil, MOPS, phenol, plastic cuvettes, Triton X-100, Tween 20

2.1.2 DNA/RNA modifying enzymes

Invitrogen™ life technologies: *Taq* DNA Polymerase, Hind III

New England BioLabs Inc.: restriction endonucleases

Roche: Klenow polymerase, M-MuLV reverse transcriptase

2.1.3 Radioactive reagents

Amersham Biosciences: Redivue [α-^{32}P]-dCTP (~3000Ci/mmole, 10mCi/ml.)

2.1.4 Antibiotics

Sigma Chemical Co. Ltd.: ampicillin, penicillin G, streptomycin

2.1.5 Mammalian cell culture reagents

Difco Laboratories: trypsin

Greiner bio-one: plasticware

Invitrogen™ life technologies: Dulbecco's modified Eagle Medium (DMEM), foetal calf serum, Glasgow Modified Eagle Medium (GMEM), L-glutamine, non-essential amino acids, sodium pyruvate

Sigma Chemical Co. Ltd.: epidermal growth factor (EGF)

2.1.6 Bacterial culture reagents

Fisher: bacto-tryptone

ICN Biomedicals/pharmaceuticals: bacto-yeast extract

Invitrogen™ life technologies: terrific broth

2.1.7 Oligonucleotides

Oligonucleotides were synthesised by the Oswel DNA service. The sequences of the primers used are provided in Table 2.1.

Table 2.1 Oligonucleotides

Table shows primer name, reference code, sequence and description. Positional numbering as in figure 3.1.

Primer name	Code	Sequence (5'- 3')	Description
2	H4395	CATGCCATGGACCCT GGGAAGGACGAG	*Ercc1* cDNA 5' end (-6 to 21)
5	F15849	TGGATGTAGTCTGGA TGGAGGTTGTGG	*Ercc1* cDNA 3' end (432 to 459)
A	F14820	TCATCACTGAGCCGG ATCTGGAGTCTGG	*Ercc1* flanking region (-317 to -290)
B	F14821	CCCAAACCACAGCG GTCCTCCAGGACC	*Ercc1* flanking region (-208 to -181)

2.2 Media

2.2.1 Bacterial culture media

Luria Broth (LB): 1% (w/v) bacto-tryptone, 0.5% (w/v) bacto-yeast extract and 0.5% (w/v) NaCl.

Luria Agar: 1.5% agar, 1% (w/v) bacto-tryptone, 0.5% (w/v) bacto-yeast extract and 0.5% (w/v) NaCl.

Terrific Broth (TB): 1.2% (w/v) bacto-tryptone, 2.4% (w/v) bacto-yeast extract, 0.4% (v/v) glycerol, 170mM potassium dihydrogen orthophosphate, and 72mM dipotassium hydrogen orthophosphate.

When required, ampicillin was added immediately prior to pouring of dishes, to a final concentration of 100μg/ml.

2.2.2 Mammalian tissue culture media

Dulbecco's modified Eagle Medium (DMEM) was used as the base culture medium for murine keratinocyte cell lines.

Glasgow Modified Eagle's Medium (GMEM) (McPherson, 1962, with modifications by W. House, Medical Research Council Institute of Virology, University of Glasgow, 1964) was used as the base culture medium for murine fibroblast and Chinese hamster ovary (CHO) cell lines.

Freezing medium for long term storage of cells consisted of the appropriate supplemented culture medium with additional DMSO and FCS to final concentrations of 10% (v/v) and 12% (v/v) respectively.

Chinese hamster ovary cell culture medium: 6% (v/v) FCS, 1X non-essential amino acids, 1mM sodium pyruvate, 25U/ml penicillin, 25µg/ml streptomycin, in GMEM.

Fibroblast culture medium: 10% (v/v) FCS, 1X non-essential amino acids, 1mM sodium pyruvate, 25U/ml penicillin, 25µg/ml streptomycin, in GMEM

Keratinocyte culture medium: 8% (v/v) chelex-treated FCS, 25U/ml penicillin, 25µg/ml streptomycin, 2mM L-glutamine, 10ng/ml EGF.

Fibroblast serum starvation medium: 0.5% (v/v) FCS, 1X non-essential amino acids, 1mM sodium pyruvate, 25U/ml penicillin, 25µg/ml streptomycin, in GMEM

2.3 Solutions and buffers

Carnoy fixative: 3 parts methanol, 1 part glacial acetic acid

Grams crystal violet: 10% (v/v) Grams crystal violet solution

Formalin: 4% (v/v) formaldehyde in phosphate buffered saline.

Formamide sample buffer (FSB): 2.4X MOPS, 48% de-ionised formamide, and 29% formaldehyde

10X MOPS – 1 litre: 200mM 4-morpholinepropanesulfonic acid (MOPS), 50mM sodium acetate, 10mM EDTA, pH 7.0

Oligo labelling buffer: 50µl solution A, 125µl solution B, 75µl solution C

Solution A: 1.25M Tris-HCl pH 8.0, 0.125M $MgCl_2$, 25mM β-mercaptoethanol, 0.5mM each of dGTP, dATP and dTTP.

Solution B: 2M HEPES buffer adjusted to pH 6.6 with NaOH.

Solution C: Random hexadeoxyribonucleotides, OD_{260nm}= 90 units/ml, in TE

Phenol/trisodium citrate EDTA (STE): STE: 50% (v/v) STE, 50% (v/v) phenol

Pre-hybridisation buffer: 6X SSC, 1% SDS, 10% (w/v) dextran sulphate, 100mg/ml denatured herring sperm DNA.

PI: 50mM Tris-HCl pH 8.0, 10mM EDTA, 100µg/ml Rnase A

PII: 20mM NaOH, 1% (w/v) SDS

PIII: 2.5M potassium acetate pH 4.8

10X PCR buffer: 500mM KCl, 1.5mM $MgCl_2$, 10mM Tris-HCl pH 8.3, 0.01% (w/v) gelatin, 0.45% (v/v) Triton X-100, 0.45% (v/v) Tween 20

QBT: 750mM NaCl pH 7.0, 50mM MOPS, 15% (v/v) ethanol, 0.15% (v/v) Triton X-100

QC: 1M NaCl pH 7.0, 50mM MOPS, 15% (v/v) ethanol

QF: 1.25M NaCl pH 8.2, 50mM MOPS, 15% (v/v) ethanol

10X RT buffer: 500mM Tris-HCl pH 8.3, 60mM $MgCl_2$, 400mM KCl, 10mM DTT

5X sample buffer: 20% (v/v) glycerol, 100mM EDTA, 0.1% (w/v) bromophenol blue

20X SSC: 3M NaCl, 0.3M tri-sodium citrate pH 7.0

STE: 10mM NaCl, 10mM tri-sodium citrate, 1mM EDTA.

10X TBE: 0.9M Tris-HCl, 0.9M boric acid, 20mM EDTA.

TE buffer: 10mM Tris-HCl, 1mM EDTA pH 8.0

2.4 Bacterial strains

The *E. coli* strain DH5α was used for propagation of plasmids. The genotype of DH5α is *supE44*, *lacU169* (ø *80lacZM15*), *hsdR17*, *RecA1*, *endA1*, *gyrA96*, *thi-1*, and was first described by Hanahan (1983).

2.5 Mammalian cell culture lines

Table 2.2 Mammalian cell culture lines

CELL LINE	DESCRIPTION	REFERENCE
PF20	Immortalised wild type mouse embryonic fibroblast cell line	Melton *et al.*, 1998
PF24^{+}#2D-4	Immortalised *Ercc1* and *HPRT* deficient murine embryonic fibroblast cell line	Melton *et al*, 1998
CHO9	Immortalised wild type Chinese hamster ovary cell line, subclone of the CHO-KK cell line	Wood and Burki, 1982
CHO43.3B (TG#1)	Immortalised *Ercc1* and *HPRT*-deficient Chinese hamster ovary cell line derived from the parental CHO9 cell line.	Wood and Burki, 1982
Keratinocyte (+/+)	Mouse wild type keratinocytes	Melton, unpublished data
Keratinocyte (-/-)	*Ercc1* and *HPRT*-deficient mouse keratinocytes	Melton, unpublished data

2.6 Methods

2.6.1 Growth of *E. coli*

Cells were grown at 37°C either in suspension in LB or TB with shaking, or on the surface of Luria agar plates. Cells grown on Luria agar plates supplemented with ampicillin were incubated at 30° to prevent growth of satellite colonies.

2.6.2 Preparation of competent cells for transformation

Bacteria were transformed with plasmid DNA by the method of Mandel and Higa (1970) with the modifications by Dagert and Ehrlich (1974). An inoculating loop was used to scrape bacterial growth from a fresh Luria agar plate of *E. coli.* and this was added to 50ml of LB supplemented with $MgCl_2$ to a final concentration of 20mM. Cells were grown with vigorous shaking at 37°C until the OD_{600nm} of the suspension equalled 0.2 (approximately 1.5-2 hours) and the cells were in the log growth phase. The cells were chilled on ice for 5 minutes, and then spun down at 2,000rpm at 4°C for 15 minutes. The pellet was gently resuspended in 20ml ice-cold freshly prepared transformation buffer. The suspension of competent cells was then incubated on ice for 30 minutes with occasional agitation and then spun for a further 15 minutes at 4°C. The supernatant was then poured off and the pellet was resuspended in 2ml ice-cold transformation buffer. The resuspended cells were then transferred to a 15ml centrifuge tube and stored on ice for 2 hours prior to transformation.

Approximately 10ng of plasmid DNA in a volume of 1µl was added to 100µl of competent cells, which were then incubated on ice for 20 minutes. Following heat shock at 37°C for 5 minutes, 400µl of warm LB was added, mixed gently, and the mixture was incubated at 37°C for ~45 minutes to allow expression of antibiotic resistance. Aliquots of 100µl and 150µl of the mixture were then spread to dryness on selective ampicillin plates (which had been dried and warmed in the incubator) and incubated overnight at 37°C. A no-DNA control was also routinely cultured.

2.6.3 Growth of mammalian CHO cells

Mammalian CHO cells were maintained in supplemented GMEM at 37°C in a 5% CO_2 atmosphere in a humidified tissue culture incubator (Sanyo CO_2 incubator). To passage cells the medium was aspirated, the bottom of the flask or dish rinsed with 1-2ml of trypsin (0.05% w/v) to remove remaining medium, then a further 1-2ml of trypsin added and incubated at 37°C until the cells had become detached. The cells were then suspended in 5 volumes of medium and transferred to a sterile 50ml tube, and then pelleted by 5 minutes centrifugation at 1,300rpm. The supernatant was discarded and the cells resuspended in an appropriate volume of medium, and an aliquot transferred to a tissue culture flask.

For long term storage, cells were kept in liquid nitrogen. Cultures to be frozen were trypsinised, spun down, the supernatant removed, and cells resuspended in ice cold freezing medium. 1ml aliquots were placed at –20°C for 2 hours then transferred to –70°C overnight before removal to liquid nitrogen. Frozen aliquots were thawed rapidly at 37°C and diluted in 9ml of medium before being pelleted by centrifugation at 1,300rpm for 5 minutes. The supernatant was discarded, the cells resuspended in an appropriate volume of medium and transferred into a flask.

2.6.4 Growth of mammalian fibroblasts

Mammalian fibroblasts were maintained in supplemented GMEM at 37°C in a 5% CO_2 atmosphere in a humidified tissue culture incubator (Sanyo CO_2 incubator). Passage and storage was performed in the same manner as in the CHO cells.

2.6.5 Growth of mammalian keratinocytes

Mammalian keratinocytes were maintained in supplemented DMEM at 37°C in a 5% CO_2 atmosphere in a humidified tissue culture incubator (Sanyo CO_2 incubator). To passage cells the medium was aspirated, the bottom of the flask or dish rinsed with 1-2ml of trypsin (0.05% w/v) to remove remaining medium, then a further 1-2ml of trypsin added and incubated at 37°C until the cells had become detached. This process takes 20-30 minutes for keratinocytes and very gentle scraping was performed if the cells did not start to detach after 20 minutes to avoid damage to cells from longer exposure to the trypsin. To remove calcium from the FCS, 0.05g of chelex was added per ml FCS and mixed for 1 hour. The solution was then filtered

through Whatman no. 1 filter paper to remove the particulate matter and then filtered through a 0.2 millipore nalgene filter to sterilise.

2.6.6 Electroporation

In order to achieve transient expression of cloned DNA sequences in cultured mammalian cells, a linearised *Ercc1* minigene vector was introduced by electroporation with a selectable *HPRT* minigene construct under the control of the mouse phosphoglycerate kinase (*PGK-1*) gene promoter. Cells to be electroporated were trypsinised, pelleted by centrifugation, and resuspended in 10ml of medium. The cells were counted using a 1:100 dilution in isoton and $5.0X10^6$ cells were then resuspended in 0.8ml of electroporation buffer (HPS) containing 120µg of plasmid DNA in 100µl sterile distilled water and then transferred to an electroporation cuvette (0.4cm, Bio-Rad). The cuvette was slotted into a Gene Pulser (Bio-Rad), and the cells pulsed at 300V, 500µF capacitance. After 10 minutes of incubation at room temperature the electroporated cells were resuspended in the appropriate volume of medium before being transferred to petri-dishes. Following overnight culture the medium was changed to selective medium containing HAT. In cases where UV selection was also required, the non-selective medium was aspirated and 0.2ml of sterile PBS was applied. Fibroblasts and CHO cells then received a dose of 7.5 Jm^{-2} of UV (254nm), the PBS was aspirated and HAT selective medium applied. The UV selection regime in keratinocytes was different because fewer cells survive to form colonies, meaning that to impose both HAT selection and UV selection can result in too few cells surviving. Therefore, HAT selection was the same as in the other cell types but a dose of 3 Jm^{-2} of UV was applied instead of 7.5 Jm^{-2}.

2.6.7 Clonal selection

Following electroporation, clones were selected and isolated using cloning rings (5mm and 10mm in diameter). Following aspiration of media the petri dish containing the clone(s) was rinsed with 1XPBS. Depending upon the size of the clone, the appropriate cloning ring was selected using sterile tweezers, dipped lightly in Vaseline to coat the underside of the ring and applied to the surface of the dish. The clone had to be completely encompassed by the ring and the clone also had to be discrete from other clones. Trypsin (0.05% w/v) was applied (1-2 drops depending

upon the ring selected). If the non-selected cells were required then 1XPBS was applied to the dish during trypsinisation to prevent drying. Once trypsinisation was complete the cells were removed from inside the ring and placed in a 5ml flask containing the appropriate growth media.

2.6.8 UV survival assays

To measure UV survival of cultured fibroblasts, 10^4 PF20 cells were plated per 30mm dish and $2X10^4$ PF24 cells were plated (the wild type cells grow more quickly so a lower plating density was used). With CHO cells, $1.5X10^4$ cells were plated per 30mm petri dish. The keratinocyte plating density was $4X10^4$ cells per 30mm dish. The following day the medium was aspirated from the dishes and replaced with 0.2ml of sterile PBS and the dishes were UV-C irradiated (254nm) at a fluence of 0.15 $Jm^{-2}s^{-1}$. Because the wild type keratinocytes are very resistant to UV, the UV lamp was lowered to a height where the fluence was 0.625 $Jm^{-2}s^{-1}$, to prevent the cells from drying out during irradiation. The UV survival assay of the transformed keratinocytes was performed with the lamp at the height that delivers a fluence of 0.15 $Jm^{-2}s^{-1}$. Following irradiation, the medium was replaced and the dishes were incubated for a further five days (or until the no-UV controls became confluent).. The cells were then fixed in Carnoy fixative and stained with Crystal Violet then gently rinsed in water prior to drying overnight. Dead cells detach from the surface of the plate during incubation and are removed during medium changes and rinsing. The degree of cell growth was then determined by extracting the dye from the stained cells using 70% ethanol and measuring the optical density (OD) at 575nm on a Thermospectronic Bioplate 3 spectrophotometer using plastic disposable cuvettes. Each UV dosage was performed in duplicate and survival was determined relative to unirradiated control dishes (representing 100% on survival index). The survival index was defined as the ratio of the mean OD of irradiated cells/mean OD of cells in unirradiated dishes. The survival index percentage was then plotted on a $\log_{10}$ scale against UV dose.

2.6.9 Preparation of plasmid DNA (Qiagen prep).

For preparation of up to 500μg of plasmid DNA, a column purification method was employed. Bacteria were grown for 24 hours at 37°C in 50ml of TB supplemented

with ampicillin to a final concentration of 100μg/ml. Cells were pelleted by spinning at 3,900rpm for 20 minutes at 4°C. The pellet was resuspended in 10ml of solution P1 supplemented with Rnase to a final concentration of 0.5μg/ml, and the cells lysed by addition of 10ml of solution P2, with 5 minutes incubation at room temperature. 10ml of cold solution P3 was added, with gentle mixing, and the tube incubated on ice for a further 20 minutes. Precipitated chromosomal DNA, SDS and protein were sedimented by centrifugation at 17,000rpm for 30 minutes at 4°C. The supernatant was removed promptly, filtered and applied to a Qiagen-tip 500 column, previously equilibrated with 10ml of QBT buffer. After washing twice with 30ml QC buffer, the plasmid DNA was eluted from the column with 15ml QF buffer. 12ml of isopropanol was added to precipitate the DNA, which was then spun down at 4,000rpm for 20 minutes at 4°C. The DNA pellet was resuspended in 500ml of sterile distilled water, prior to transfer into a microfuge tube, whereupon DNA was reprecipitated with 50μl of P3 and 1ml of cold absolute ethanol. Following 10 minutes centrifugation, the nucleic acid pellet was washed twice with 70% (v/v) ethanol, air dried, resuspended in 200μl sterile distilled water and stored at –20°C.

2.6.10 Purification of plasmid DNA by phenol extraction

DNA (usually diluted in 400μl) was added to an equal volume of phenol/STE in an Eppendorf tube, vortexed and centrifuged in a microfuge until the mixture separated into two layers. The top layer was transferred to a fresh tube and half of this volume of butanol was added. This mixture was vortexed and centrifuged in a microfuge (Hettich MIKRO 20) at 13,000rpm for 1 minute or until the mixture separated into two layers. The top layer was then discarded and one tenth of the volume of 3M sodium acetate (pH 5) was then added. The mixture was vortexed and 2.5X the volume of chilled absolute alcohol (EtOH) was added to precipitate the DNA. The mixture was vortexed and centrifuged for 7 minutes or until the DNA pellet formed. The supernatant was then carefully poured off and the pellet was rinsed with a small volume of 70% EtOH. The EtOH was gently poured off and the pellet was air-dried then resuspended in the desired volume of distilled H_20 prior to OD measurement.

2.6.11 Preparation of mammalian RNA (RNA-Bee™ method).

A modification of the method described by Chomczynski and Sacchi (1987) was used for the isolation of total RNA from all cultured cells and tissues. The method used the commercially available reagent RNA-Bee™ (Biogenesis) and followed the protocol supplied by the manufacturers. Briefly, tissue was homogenised with 1ml RNA-Bee™/50mg tissue using an Ultra-Turrax T25 homogeniser (IKA-Laboratechnik). Cultured cells were lysed directly in the culture dish by the addition of RNA-Bee™ (1ml per 3.5ml petri dish) and the RNA was solubilised by pipetting the lysate a few times. The RNA was extracted by the addition of 0.2ml chloroform per 1ml of RNA-Bee™ used, followed by vigorous shaking of the mixture for 15-30 seconds prior to placing on ice for 5 minutes. The suspension was then centrifuged at 12,000g for 15 minutes at 4°C. The aqueous phase was transferred to a fresh tube and the RNA precipitated by the addition of an equal volume of isopropanol. The RNA was left to precipitate at 4°C for 5-10 minutes and then pelleted by centrifugation at 12,000g for 5 minutes at 4°C then washed in 75% ethanol (using at least 1ml of ethanol per 1ml of RNA-Bee™ used for the initial homogenisation)

2.6.12 Estimation of DNA concentrations via OD measurement

DNA samples were diluted in 1ml of distilled water and the OD at 260nm and 280nm measured in a spectrophotometer (Thermospectronic Bioplate3) using matched quartz cuvettes. 1 A_{260} Unit of dsDNA = 50µg/ml. The ratio of A_{260}/A_{280} gives an estimate of nucleic acid purity. A value ≥ 1.8 indicates that the sample is pure DNA. A value below 1.8 indicates that the preparation is contaminated with proteins and aromatic substances while a value above 2 indicates possible contamination with RNA.

2.6.13 Estimation of RNA concentrations via OD measurement

RNA samples were diluted in 1ml of distilled water and the OD at 260nm and 280nm measured in a spectrophotometer (Thermospectronic Bioplate3) using matched quartz cuvettes. 1 A_{260} Unit of ssRNA = 40µg/ml H_2O. The ratio of A_{260}/A_{280} gives an estimate of nucleic acid purity. A value ≥ 2.0 indicates that the sample is pure RNA. A value below 2.0 indicates that the preparation is contaminated with proteins and aromatic substances.

2.6.14 Restriction of DNA with endonucleases

See figures 3.1, 3.2, 3.3 and 3.4. DNA digests were performed to ensure the identity of minigene DNA following Qiagen preps using a single *EcoRI* digest and a double *kpn1/EcoRI* digest. Reactions were carried out in accordance with temperature and buffer conditions recommended by the manufacturer. The double digest was performed in two stages; the *kpn1* digestion was performed in React 1 buffer for 1.5-2 hours prior to addition of *EcoRI* and React 3 buffer for further digestion. The *Kpn1* and *EcoRI/kpn1* digests were terminated by the addition of 5X sample buffer for agarose gel electrophoresis. Once the *kpn1/EcoRI* and *EcoRI* digestion had verified the minigene identity, the minigenes could be linearised with *Not1*. 1% agarose gel electrophoresis was performed to ensure minigenes were fully digested by *Not1* prior to electroporation. A second round of *EcoRI* and *kpn1/EcoRI* restriction was performed to double-check the identity of the minigenes following *Not1* linearisation.

2.6.15 Electrophoresis of DNA in agarose gels

DNA fragments were separated on 1-2% (w/v) electrophoresis grade agarose gels containing 0.5µg/ml ethidium bromide and 1X TBE. DNA samples were mixed with one-fifth volume of 5X sample buffer prior to loading. Electrophoresis was carried out horizontally at 100-140V in a 1X TBE buffer system.

Hyperladder I™ was used as a size marker. This ladder contains 14 regularly spaced bands ranging from 200bp to 10Kb, the 1Kb and 10Kb bands have the highest intensity. Following electrophoresis, DNA/RNA was visualised in a High Resolution UV Gel Documentation system and digitally photographed where necessary. Where necessary, gel photographs were analysed using the public domain NIH Image program, ImageJ software (version 1.3), developed at the U.S. National Institutes of Health and available on the Internet at http://rsb.info.nih.gov/nih-image/.

2.6.16 Northern blot - electrophoresis of RNA in agarose gels

RNA samples were electrophoresed on denaturing 1.4% (w/v) ultraPURE™ agarose containing 0.5µg/ml ethidium bromide, 1X MOPS and 0.66M de-ionised formaldehyde. Formaldehyde and ethidium bromide were added to the melted

agarose mixture following cooling to 75°C and then gels were poured and set in a fume-hood.

Samples were prepared by the addition of 30µg of total RNA (in 20µl dH_2O) to an equal volume of FSB. This mixture was heated for 5 minutes at 65°C, snap-cooled on ice and one quarter volume of 5X sample buffer was added. The samples were then loaded directly. Gel resolution was improved by performing electrophoresis at 4°C (in a refrigerated room). Electrophoresis was performed at 100V for 3-6 hours in a 1X MOPS buffer system.

2.6.17 Northern blot - transfer of RNA from agarose gels to membranes

Following electrophoresis the gel was soaked for two 20 minute periods in 10X SSC (a dilution of 20X SSC buffer), with gentle agitation to remove ethidium bromide and formaldehyde. The Genescreen *Plus*®nylon membrane was soaked for 5 minutes in 10XSSC at room temperature. RNA was transferred to the Genescreen membrane by capillary action using 10XSSC as the transfer medium. When transfer was complete (12 to 24 hours) the Genescreen membrane was marked with pencil to indicate the side that was in contact with the gel and rinsed in 2XSSC for 10 minutes with gentle agitation, air dried and then baked for 2 hours at 80° C under a vacuum (Yallenkamp vacuum oven).

2.6.18 Northern blot - randomly primed labelling method

An 870bp restriction fragment from the *ERCC1* cDNA clone pTZME (kindly provided by M. van Duin (described in van Duin *et. al.*, 1988) was used detect *ERCC1* mRNA. This fragment extends from the 5′ end to a *BamHI* site in exon 8. Radioactive labelling of DNA to high specific activity was performed by the randomly primed labelling method (Feinberg and Vogelstein, 1983). DNA probes were denatured, annealed to random hexamer primers, part of the oligolabelling buffer (OLB) solution, and the complementary strand was synthesised from the 3′OH termini of the hexanucleotide primers using Klenow polymerase (the large fragment of DNA polymerase I) in the presence of [α^{32}P]dCTP (3000Ci/mmol). The radiolabelled dCTP is incorporated into the newly synthesised DNA strand. These

radiolabelled DNA fragments serve as efficient probes in filter hybridisation. 25ng of DNA, dissolved in 32µl of distilled, de-ionised H_2O, was denatured by boiling for 5 minutes and snap chilled on ice. 10µl of OLB, 20µg of BSA, 50µCi of [α^{32}P]dCTP and 7.5 units of Klenow polymerase were added, and the mixture was incubated for 3 hours (or overnight) at room temperature [final volume of mixture =50µl]. 2µl of this mixture was spotted onto Whatman GF/C glass microfibre filters (2.1cm diameter) for measurement of [α^{32}P]dCTP incorporation. Prior to hybridisation, 2.25mg of sonicated herring sperm DNA was added to the probe, which was then denatured by boiling for 5 min.

2.6.19 Northern blot - separation of unincorporated radionucleotides

The probe was purified with TE buffer through a Nick™ Column (Amersham Biosciences) according to the manufacturer's instructions. Briefly, the column is first equilibrated with TE buffer and once the equilibration buffer has completely entered the gel bed the probe can be added. Once the probe has entered the gel bed 400µl of TE is added and allowed to drain through the column into an Eppendorf tube. The column is now placed onto a new Eppendorf tube and another 400µl of TE is allowed to drain through. This elution removes the probe from the column while unincorporated radioactivity remains in the column. 8µl of the purified probe mixture is spotted onto 2 glass microfibre filters (4µl/filter). The filters were placed in scintillation vials at this point and put aside to dry, prior to TCA precipitation.

2.6.20 Northern blot - measurement of radionucleotide incorporation by TCA precipitation

To calculate the proportion of [α^{32}P]dCTP incorporated into the probe, differential precipitation by trichloroacetic acid (TCA) was performed. The glass filters spotted during radioactive labelling were in two pairs, one pair was spotted prior to nick column purification (and therefore contains both incorporated and unincorporated radionucleotides) while the other pair of filters were spotted following purification. In each pair of filters, one filter was used to measure the total amount of radioactivity in the reaction and the other was used to measure only the acid-precipitable radioactivity. DNA molecules more than 50 nucleotides long are precipitated onto the surface of the filter. To perform precipitation, one filter from each pair was

transferred (using forceps) to 50ml capped tubes containing 45ml of ice-cold 5%(w/v) TCA. The filters were gently agitated for 2 minutes and then transferred to fresh tubes containing the same volume of ice-cold TCA. The washing was repeated two more times. The washed filters were then transferred to fresh tubes containing 70% ethanol, left briefly, then dried at room temperature. The four filters (two washed and two unwashed) were inserted into individual scintillation vials. Following addition of scintillation fluid, the ^{32}P radioactive decay was measured in the ^{32}P channel of the liquid scintillation counter. These values were used to calculate the proportion of radionucleotide incorporated, the total amount of product and the specific activity.

2.6.21 Northern blot - pre-hybridisation and hybridisation

Pre-hybridisation of the genescreen membrane was performed in a temperature-controlled rotisserie in 30ml of prehybridisation buffer at 60°C for a minimum of 2 hours. 15ml of pre-hybridisation buffer was removed prior to addition of the probe. The probe was denatured by boiling for 5 minutes in a screw-cap tube with 225µg of 10mg/ml sonicated herring sperm DNA (2.25mg) and promptly added to the pre-hybridisation solution with the glass vial in a vertical position to avoid dripping concentrated probe on the genescreen membrane. Hybridisation was performed overnight at 60°C. After hybridisation, the solution was discarded and the blot was rinsed for 2X5 minutes in 2XSSC at room temperature, 2X30 minutes in 2XSSC, 1% (w/v) SDS at 65°C and 2X30 minutes in 0.1XSSC at room temperature. The blot was then sealed in cling-film and the activity was measured with a Geiger counter to estimate the optimal exposure time.

2.6.22 Autoradiography

Blots were autoradiographed using Kodak X-ray film in a cassette with intensifying screens (Cronex Lightning Plus, Du Pont) and stored at –70°C to slow reversion of activated bromide crystals in the film to their stable form. The resulting signals were analysed using the public domain NIH Image program, ImageJ software (version 1.3), developed at the U.S. National Institutes of Health and available on the Internet at http://rsb.info.nih.gov/nih-image/.

2.6.23 RT-PCR - reverse transcription

Reverse transcription was carried out on 20μg of total RNA. To denature the RNA, samples were heated to 70°C for 2 minutes and then snap cooled on ice. Reverse transcription was carried out at 43°C for 1 hour in 1X RT-PCR buffer with 0.5μl Rnasin, dNTPs to a final concentration of 1.25mM, 0.1μg of random hexamers, 25μg of oligo-dT, and 30 units of reverse transcriptase. The reaction was then heated at 75°C for 1 minute and then snap cooled on ice. A second round of reverse transcription was carried out by the addition of a further 0.5μl Rnasin and 30 units of reverse transcriptase. After incubation at 43°C for 1 hour the reaction was diluted in 200μl of TE and stored at 4°C. For subsequent procedures 10μl of the total cDNA pool was used.

2.6.24 RT-PCR - amplification of DNA by the polymerase chain reaction.

PCR was carried out in 50μl volumes in 1XPCR buffer supplemented with EDTA and dNTPs to final concentrations of 0.4mM and 0.1mM respectively. 2.5 units of *Taq* DNA polymerase, 10μl of the total cDNA pool from the RT reaction plus 300ng of each primer were used per PCR. Cycle conditions were 94°C for 1 minute to denature, 68°C for 1 minute to anneal and 1.5 min synthesis with 28 cycles for the A/5 and 2/5 primer pairs and 30 cycles for the B/5 primer pair. Reactions were performed on a Techne PHC-2 PCR machine.

2.6.25 Animal husbandry

Mice were maintained in accordance with established animal care guidelines (Poole, 1989), and all procedures were licensed by the Home Office under the Animals (Scientific Procedures) Act 1986. The mice were kept under a 12 hour light and 12 hour dark cycle, with standard mouse diet and tap water. The mouse strain was C57BL/6.

2.6.26 Animal procedures - isolation of primary fibroblasts

The method of primary fibroblast isolation varied depending upon the age of the embryo or pup. Sacrifice of the pregnant mice (to obtain embryos) and pups, for analysis of *ERCC1* transcript expression, was performed by trained members of staff in our laboratory. Fibroblast isolation was performed using sterile instruments in a

primary culture hood. In experiments where light exposure had to be controlled, the stages following decapitation were performed in low light levels and the cells were shielded from light during incubation. Every effort was made to minimise light exposure. Sacrificed embryos/pups were shielded from light, all fluorescent lights of the hood and work area were turned off and window blinds were adjusted so that the ambient light level was very low (only enough visibility to work by). Tissue not being manipulated was kept covered with aluminium foil. The only stage that needed more illumination was the separation of dermis from epidermis and this was performed as quickly as possible to minimise light exposure. In light/dark experiments, a dish was designated the 'viewing dish' and this dish was synchronously passaged with the 'dark' dishes to track density during the cell expansion stage, reducing exposure of the 'dark' dishes to microscope light.

2.6.27 Isolation of primary embryonic fibroblasts from 13.5-day-old embryos

The murine uterus was transferred to 50ml tube containing 37°C fibroblast medium. The tube was transferred to the cell culture suite in a vacuum flask containing water warmed to 37°C. The tube was wrapped in aluminium foil for the light/dark experiments.

The embryos were decapitated using a scalpel and dipped in ethanol (100%) to sterilise, then rinsed in sterile PBS.

Because the skin was too delicate to be removed manually at this age, each embryo was placed into a 50ml tube containing 0.5ml of trypsin (0.05% w/v) and incubated at 37°C for 20-30 minutes.

The trypsin was then aspirated and the embryo was rinsed in sterile PBS, which was then aspirated. 15ml of fibroblast medium was then added to each tube. Each embryo was pipetted ~10X to break up the embryo prior to plating.

Cells were plated in the appropriate volume of fibroblast medium in 90mm petri dishes and incubated in the normal conditions for fibroblast culture (see fibroblast culture method). During the primary culture experiments the cells were not permitted to reach confluence prior to RNA sampling.

2.6.28 Isolation of primary embryonic fibroblasts from 16.5-day-old embryos

The method of fibroblast isolation from 16.5-day-old embryos is the same as with 13.5-day-old embryos up to the trypsin incubation stage.

Following sterilisation and rinsing in sterile PBS, each embryo was placed in a 50ml tube containing 1.5ml of trypsin (0.05% w/v) and incubated at 37°C for 1 hour.

The trypsin was then aspirated and the embryo was rinsed with sterile PBS, which was then aspirated. The skin was then removed (as described for 18.5-day-old embryo section) and chopped up using scalpels before being placed in a 50ml tube containing 5ml of fibroblast medium. The skin was then pipetted and the tube was shaken vigorously to separate the cells.

The cells were then plated in 60mm petri dishes with the appropriate volume of fibroblast medium (with the equivalent of no more than 3 skins per dish to avoid peeling).

2.6.29 Isolation of primary fibroblasts from 18.5-day-old embryos and pups.

Embryos older than 16.5 days and pups were sacrificed and rinsed in sterile PBS. The limbs and tail were removed from the embryos/pups and a slit was made down the spine that extended through the epidermis and dermis but not into the underlying tissue. The edge of the skin at the slit was then teased from underlying tissue using tweezers and, gripping the underlying tissue with blunt-ended tweezers, the skin was gently detached from the underlying tissue by pulling and rolling around the second pair of tweezers, taking care not to tear the skin.

The skin was then rinsed in ethanol (100%) and then sterile PBS before being floated dermis side down in a petri dish containing trypsin (0.05% w/v). The skin was left for 12 hours at 4°C (must not be left for longer or it becomes difficult to separate the skin layers without tearing).

Following trypsinisation, the dermis and epidermis were teased apart using tweezers. The epidermal layer is extremely delicate and must be slowly and gently pulled away from the dermal layer.

Once the two layers had been separated, the dermal layer was sliced using tweezers and added to a 50ml tube containing 5ml of fibroblast medium. The dermis was then shaken vigorously/pipetted to separate the cells. The mixture was passed through a fine sieve and suspended in the appropriate volume of medium prior to plating out in 60mm petri dishes. No more than the equivalent of 3 skins was plated per 60mm petri dish, as this is sufficient to create a monolayer of cells. If the cells are plated too thickly the cells adhere to the bottom of the dish and then peel away. Incubation conditions are described in the fibroblast culture method.

2.6.30 UV-irradiation of primary cultures

Dermal fibroblasts from a pool of seven E18.5 skins were isolated and cultured in near darkness. Following expansion in culture for 7 days, two 90mm petri dishes of cells approaching confluence were pooled and split between eight 90mm petri dishes (passage 1). 24 hours later medium was aspirated from the dishes and sterile PBS was applied to prevent the cells from drying out during irradiation. The UV-C dishes were irradiated at a fluence of 0.15 $Jm^{-2}s^{-1}$. The dose delivered was 6 Jm^{-2} and the UV lamp was at the same height as for the UV survival assays. The UV-B irradiation was performed with the UV lamp at the same height as for the UV-C irradiations and cells were irradiated for 5 minutes. Control dishes were not UV irradiated. The 'light control' dishes were not shielded from light during the medium change. Dark control dishes were shielded from light during the medium changes.

2.6.31 Serum starvation experiment.

Dermal fibroblasts from a pool of seven E18.5 skins were isolated and cultured in near darkness. Following expansion in culture for 10 days, four 90mm petri dishes of cells approaching confluence were pooled and split between eight 90mm petri dishes (passage 2). 24 hours later, 5 dishes were aspirated and serum starvation fibroblast medium was applied. The medium on the remaining dishes was changed for fresh medium. The cultures were released from serum starvation 48 hours later and RNA was extracted post-release at 0h, 3h, 6h, 20h and 48h. RNA was extracted from the control dishes at 0h and 48 hours.

Chapter 3

Use of *Ercc1* minigenes to complement DNA repair deficiency in *Ercc1*-deficient cultured mouse embryonic fibroblasts.

3.1 Control of gene expression in eukaryotes.

Before discussing the promoter of the *Ercc1* gene, which is atypical of higher eukaryotic genes, it is necessary to present a summary of how eukaryotic promoters are usually arranged. Expression of eukaryotic genes is controlled via the interaction of a set of proteins, trans-acting transcription factors, with a region of the gene called the promoter and sequences known as promoter-proximal elements and enhancers. Transcription factors recognise and bind such motifs in the genetic sequence and are, therefore, specific to particular genes. RNA polymerase II, responsible for synthesis of mRNA, is recruited by transcription factors for transcription of genes. By this process, transcription factors initiate transcription, enabling expression of genes to be controlled and modified in response to external or internal environmental stimuli.

Promoters, promoter-proximal elements and enhancers are all types of cis-acting regulatory sequences to which the trans-acting transcription factors interact. Interaction with promoters initiates transcription while the rate of transcription is increased by the presence of the promoter-proximal elements and enhancers; as the name suggests, enhancers enhance the rate of transcription. Promoters are usually located upstream of the transcriptional start site (TSS), also known as the transcriptional initiation site (TIS). Promoter-proximal elements are obviously close to the promoter (and TSS) while the enhancer elements are often quite distal to the TSS.

Three conserved sequence elements usually located near the TSS in promoters of higher eukaryotes are the TATA, CAAT and GC boxes and these classical promoter elements are usually located approximately 30, 100 and 200 bp upstream of the TSS respectively. The consensus sequence of the TATA box is TATA (T/A)A(T/A), that of the CAAT box is GGCCAATCT and the GC box sequence is GGGCGG. The TATA box recruits the components of the RNA polymerase II initiation complex and RNA polymerase II then initiates transcription approximately 30 bp or less downstream of the TATA box (in mammalian systems), so the TATA box determines where transcription of the gene begins. TFIID forms part of the RNA polymerase initiation complex and this is composed of a number of sub-units. One of these sub-units is the TATA-box binding sub-unit (*TBP*) and it is this molecule

that initially recognises the TATA box. This is followed by recruitment of the remaining RNA polymerase II initiation complex sub-units. The CAAT and GC boxes enhance the efficiency of the TATA box. The CAAT box enhances the frequency of transcriptional initiation while the GC box (often present in multiple copies) binds with a transcription factor called Sp1. Although these boxes are found in most eukaryotic promoters there are exceptions. Some genes expressed at very low levels contain none of these three boxes but do have a CG-rich sequence upstream from the TSS, which is located nearby. Aside from the CAT and GT boxes, other upstream elements include the sequences GCCACACCC and ATGCAAAT.

The enhancers of a gene can influence the efficiency of the promoter from as great a distance as 50kb. Unlike the promoter elements, which are usually located upstream of the TSS, enhancers can be located upstream or downsteam of the promoter and still influence transcription efficiency. As mentioned earlier, there is specificity in the interaction of transcription factors with their target genes. There are transcription factors that bind to promoter regions and others that bind to enhancer regions. This is possible because enhancers contain sequence motifs that are recognised by particular transcription factors. Enhancers are important in higher eukaryotes for differential control of gene expression. Some enhancers can respond to changes in external environmental conditions such as temperature fluctuations or internal environmental changes such as the production of steroid hormones. This is very important because different gene products may be required at different times or different locations in the body depending upon the prevailing environmental conditions. A second property of enhancers is that some are capable of directing tissue-specific gene expression.

The *ERCC1* gene is unusual because it does not have the genetic motifs that are normally associated with eukaryotic gene expression. The human *ERCC1* promoter was characterised (van Duin *et al.*, 1987) and does not contain any of the classical promoter elements associated with eukaryotic genes. The promoter was found to reside within 170bp of the transcriptional start site. When this 170bp region was joined 5′ to a complete human *ERCC1* cDNA, the DNA damage repair deficiency of *Ercc1* deficient CHO43-3B cells was corrected. As the novel skin-specific *Ercc1*

transcript initiates 5′ to the 170bp region that contains the promoter of the normal *Ercc1* transcript a second, skin-specific promoter must be present. A 1.05kb 5′ flanking sequence of the *Ercc1* gene was cloned and sequenced by Kan-Tai Hsia (Figure 3.1) and utilised in construction of the minigenes used for the transfections illustrated later. This sequence contains the 170bp region that contains the normal promoter and should also contain the skin-specific promoter because of the likely proximity to the skin TSS. Sequencing this region revealed the presence of a number of interesting sequence motifs. A TATA box (consensus sequence TATATAA) was found at nt position –823 to -829 and a GC box (consensus sequence GGGCGG) was located at nt position –975 to -980. However, the TATA box is approximately 322bp upstream of the putative TSS of the skin transcript and since TATA boxes are usually within 30bp of the TSS this TATA box is probably either non-functional or coincidental. GC boxes are usually located within 200bp of the TSS so the GC box is also a little distal to be acting as a promoter-proximal element. Two interesting sequence motifs in the fragment are a stretch of CA repeats at the 5′ end and stretch of CT repeats. One of the polypyrimidine/polypurine CT regions repeats 34 times $(CT)_{34}$ while a second repeats 4 times $(CT)_4$. Counting from the A of the normal transcript TSS being designated as nt +1, the $(CT)_{34}$ repeat begins at nt –428 while the $(CT)_4$ begins at nt –358. Since the region responsible for transcription of the skin-specific transcript is only used in skin it is possible that tissue specificity is achieved by regulatory enhancer regions.

Figure 3.1 Nucleotide sequence of the murine *Ercc1* 5′ flanking region

The nucleotide sequence of the murine *Ercc1* 5′ flanking region (pRK1.05) is shown. The stretch of CA repeats is highlighted in red. Consensus sequences for GC and TATA boxes are highlighted in yellow and green respectively. The nt position for initiation of transcription in skin is shown in pink. The normal transcriptional start site is highlighted in dark green. *BseRI*, *BclI* and *kpn1* restriction sites are in red, underlined and have italic tags. The CT repeat tract is highlighted in blue. Data kindly provided by Kan-Tai Hsia. Figure adapted from original by Jim Selfridge. Nucleotide numbers are numbered with the 'A' of the normal transcript's translational start site as nt +1.

-1284 GAATTCACAG CGTAGGCATG TGCTTGTACA CACACACACA CACACACACA CACACACACA -1225

-1224 CACACACACA CACACACACC TCACAACTAG GAAAACATTA AAAAAGATAT AGAGGGATTA -1165

-1164 CATATGATGA CCTTTGGGCG TGCCCTCTCT CAATACACGA AGACCAAAAG AAAATGGAGA -1105

-1104 ATACGTTGGT AAATGCATGC AAAATTTTAA ATGGAGAATT TTTAAAAGTA TGAACTCGAA -1045

-1044 GGAAAGATTT GCAGCCTAAT GAAGTGACTA ACAGCATCAG CCCTTAAGCA TCACTCCAAC -985

-984 AAATGGGCGG TAATCAAAAG ATGCTCATCT CCTTTTCCTC CCCCTTTTTA ATTTAAAAAA -925

-924 ACCGTTGAAT TGCTCATTTC GAAAAAATGG GCAAAGGATA TGAGCCAATC TTTGTTAAAA -865

-864 GGAATAGGCA AAGGAAATGA GCAAGAACAA AAAAATATAT AAGACCATTA AACTATTTCC -805

-804 TTAGCAATCA CAGAACTACA AGTTGGAACG GGGTGAAACT GTCCTTCCAG GAAATGAGTC -745

-744 ATACTAAGTC ATTCATGGGG CCCAGGTCTC GGAACACAGG TGTCCCTTCC TTCAGGAAGC -685

-684 TCGTCCTCAC ACCCCAGCTG GGTCAGGTGC TCCCCCTAAC ACGGGCTCAA TGTCCTCTGG -625

-624 GCTCTCCCTT CCCGACCCTG CCCGCTCTGG GTTGTCCCTG TCTGACTGTC CCACTGGACT -565

-564 TTAAGCCACA GCTGTCGGTA TCAGTCCCTG CTTGGGATGC AAGTGGGGAA TGGCTTAGAA -505

-504 TTAGTGAGTC AATGTCTGTG GGCTGAGTCC ACCCACGCTT ACAGGGAAGG GCGTCCTCTG -445

BseRI

-444 AAATCTCACC TCCCTCCTCT CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT -385

-384 CTCTCTCTCT CTCTCTTTCT CTCTGTCTCT CTCTCCTTTC CCTCTTTCCC TCCCCCACCT -325

BclI

-324 CCTTTGATCA TCACTGAGCC GGATCTGGAG TCTGGGAAGC GCTCAAGAGG GCCTTGGAAC -265

kpnI

-264 ACAACTACTT GAAGTCAAAG TCTCCCAGGT ACCAAGATGC ACAAGCTTCC ATCCGCCCCA -205

-204 AACCACAGCG GTCCTCCAGG ACCATAGAGA GCAGCGCGAA TAGAGTTCCC CGCTCTAACT -145

-144 CCTCCGGGGA GCAGCGAGAC GAGCGAAGGG CCAGAGGGCC GGAAGTGAGT CTAGCAGGAG - 85

- 84 TTGTGCTGGC TGTGCTGGCG TTGTGTCGCC TCTGTTTCCC CCCGTGGTAT TTCCTTCTAG - 25

- 24 GCATCGGGAA AGACCAGGCC CCAGATG +3

Promoters can be identified by the use of reporter genes. However, this method was not suitable for identification of the skin-specific transcript promoter because of the presence of the normal promoter in the 5′ flanking region of the *Ercc1* gene. Attaching the 5′ flanking region of the murine *Ercc1* gene to a reporter gene would result in expression of the reporter gene product. However, this expression would result from interaction with both the normal promoter, 170bp upstream of the TSS, and any elements functioning as a skin-specific transcript's promoter. Even if the skin-specific transcript's promoter could be rendered non-functional by deletion of a crucial region the reporter gene would still be expressed due to the presence of the normal promoter. To be certain that the reporter gene expression resulted from transcription due to the skin-specific transcript's promoter, the normal promoter would have to be deleted from the region flanking the reporter gene. However, deletion of the normal promoter could have unexpected results. It is a possibility that the normal promoter may be required for the skin-specific promoter elements to function correctly. The cell type in which the experiment was conducted would also be important since the skin-specific transcript is obviously not expressed in all tissues.

3.1.1 Selection of cells suitable for transfection with *Ercc1* minigenes.

Due to the high levels of nucleases associated with mouse skin it is technically problematic to isolate RNA from this tissue for study of *Ercc1* transcripts. A panel of cultured cells available in the Melton laboratory was screened by northern blot and several cell lines were identified that had useful *Ercc1* expression patterns for investigation of the novel *Ercc1* transcript that had been identified in mouse skin. Six cell lines, three pairs, were selected for use in minigene transfections. A wild type mouse keratinocyte line, Ker. (+/+), expressed both *Ercc1* transcripts with a high 1.5kb to 1.1kb transcript ratio and this is the pattern of expression observed in skin; an *Ercc1* and *Hprt* null murine keratinocyte line, Ker. (-/-), was available for transfection with minigenes. Mouse embryonic fibroblasts express both the 1.5 and 1.1kb *Ercc1* transcripts at equal levels and a wild type cell line, PF20, was available; an *Ercc1* and *Hprt* null mouse embryonic fibroblast line, PF24, was also available. Finally, Chinese hamster ovary cells express only the 1.1 kb transcript and a cell line was available, CHO43.3B (TG#1); an *Ercc1* and *Hprt* null CHO line was also

available. The three *Ercc1* and *Hprt* null cell lines are hypersensitive to UV and HAT while their wild type counterparts are resistant to UV and HAT. The mouse embryonic fibroblasts (PF20) are dermal in origin and, therefore, have an expression pattern that mimics that of the in vivo dermal *Ercc1* expression pattern. The wild type keratinocyte line is epidermal in origin and this is why they have an expression pattern that mimics that of in vivo epidermis. When skin is processed as a whole tissue it includes the dermis and epidermis, so the *Ercc1* pattern of expression observed in skin is like that of keratinocytes (with a high 1.5 to 1.1kb transcript ratio). The cell lines and the *Ercc1* transcripts they produce are summarised in table 3.1 and the expression pattern was shown in figure 1.2.

Table 3.1 Cells used for *in vitro* *Ercc1* expression pattern investigations.

Table shows the *Ercc1* and *Hprt* null cell lines chosen for use in the *Ercc1* minigene transfections and their wild type counterparts. Also detailed are the *Ercc1* transcripts produced by the wild type cells.

Wild type cell line	***Ercc1* expression pattern of wild type cell line**	***Ercc1* and *Hprt* null counterpart**
Chinese Hamster Ovary cell line: CHO#9	1.1kb transcript only	CHO43.3B (TG#1)
Mouse embryonic fibroblast cell line: PF20	Equal ratio of 1.5kb to 1.1 kb transcript	Mouse embryonic fibroblast cell line: PF24
Mouse keratinocytes: Ker. (+/+)	High ratio of 1.5kb to 1.1kb transcript	Mouse keratinocytes: Ker. (-/-)

3.2 Structure of Ercc1 minigenes

Three *Ercc1* minigenes were available in the Melton laboratory. These had been constructed in order to identify regions of the 5′ flanking sequence of the *Ercc1* gene required for production of the skin-specific transcript. One of the minigenes, *Ercc1* MG#13, had been used to transform *Ercc1*-null fibroblasts, Chinese hamster ovary cells and keratinocytes. Subsequent UV survival analysis had demonstrated that this minigene was capable of correcting the UV hypersensitivity of these cell lines. The minigene was, therefore, functional and transfection resulted in production of ERCC1 protein. Deletion of regions in the flanking sequence led to production of two further minigenes for use in transfections. This could enable identification of regions required for production of the skin-specific transcript. Deletion of essential regions could result in changes in the UV hypersensitivity of the transfected cells. Loss of the skin-specific transcript following deletion of crucial regions would be identified by RT-PCR analysis of RNA extracted from transfected cells. To this end, *Ercc1*MG#11 and #20 had been constructed.

The use of minigenes is subject to a number of limitations. Because minigenes randomly integrate into the genome, total expression levels can be influenced by the number of minigenes incorporated, the copy number, and by random associations with chromosomal elements such as enhancers and silencers situated near the integration site of the minigene. It is possible for multiple copies of a minigene to successfully integrate into the genome of a particular cell and higher copy numbers may result in higher levels of transgene product being expressed than would be the case with the endogenous gene. Because of the variability in the location of minigene integration in the genome, it is possible that the resultant expression level may be enhanced or reduced depending upon the proximity to expression control elements such as enhancers or silencers (such elements are discussed in greater detail in the introduction to chapter 3). Assuming the correct minigene copy number, transgene expression could still be abnormal compared to endogenous gene expression due to interaction between the minigene and expression control elements leading to either higher or lower expression depending upon the element involved.

Stable integration of minigenes results in stable expression while failure to integrate leads to transient expression.

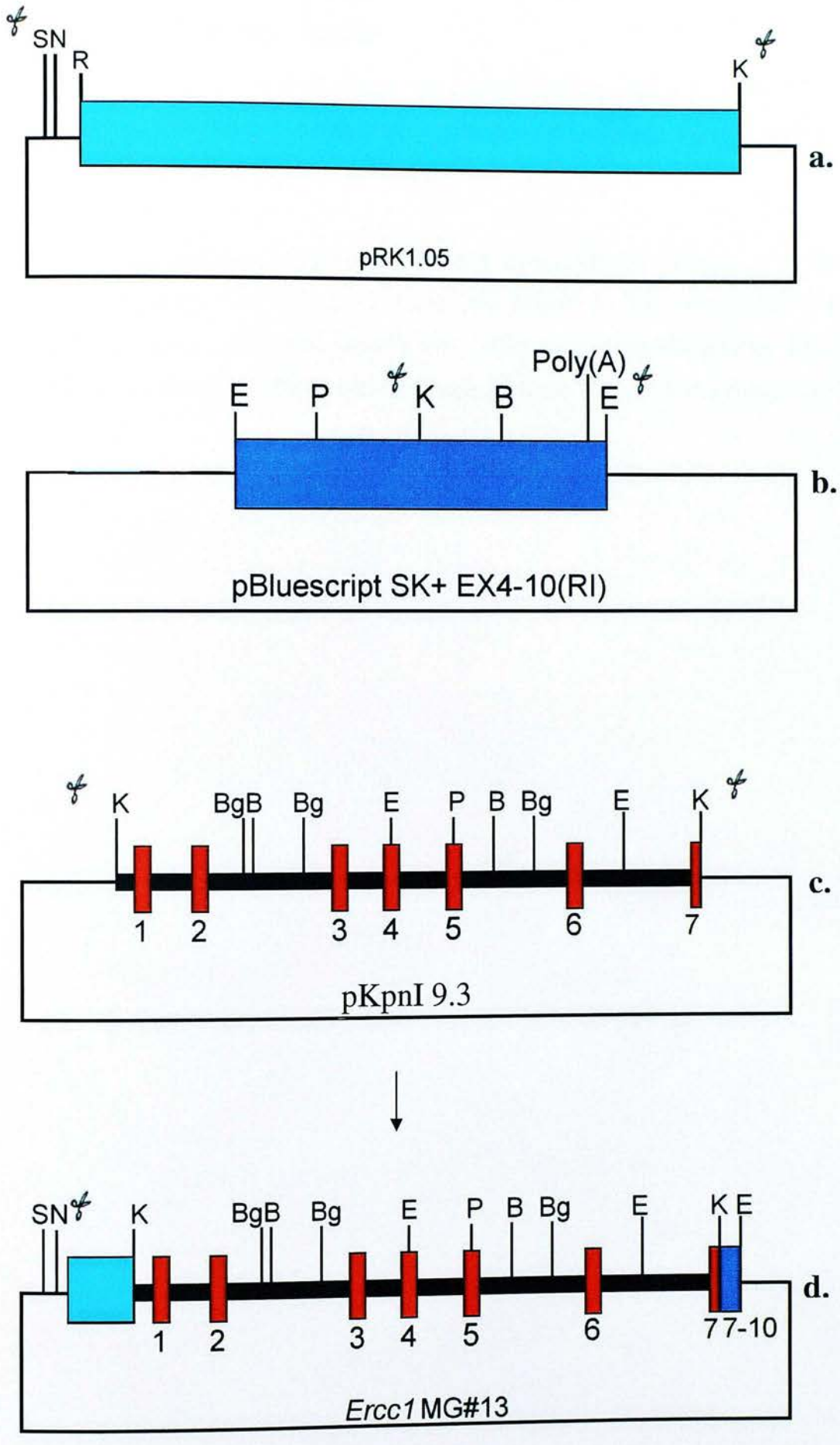
3.2.1 Construction of *Ercc1* MG#13

Ercc1 MG#13 was assembled using the cloning strategy shown schematically in Figure 3.2. The minigene was constructed by linking a 1.05kb stretch of the 5′ flanking sequence of the murine *Ercc1* gene (mentioned earlier with regard to the presence of the interesting (CT) repeat region) to an *Ercc1* gene fragment containing exons 1-10 and the 3′ end of the *Ercc1* cDNA which contains the polyadenylation signal. Construction of *Ercc1* MG#13 was performed by Jim Selfridge using three parental plasmids available in the Melton laboratory. The 5′ flanking region was isolated from the genomic subclone, pRK1.05, as a 1.05kb *SstI/KpnI* fragment. The *Ercc1* gene fragment which contained exons 1-10 was constructed by joining a 9.3kb genomic subclone pKpnI 9.3, containing exons 1-7, to a 300bp *KpnI/EcoRI* fragment of the 3′ end of the *Ercc1* cDNA containing exons 7-10, the polyadenylation signal and a short 3′ UTR. The *KpnI* site in exon 7 was utilised in the joining of these two components. The assemblage was cloned into *SstI/EcoRI* cut pBluescriptII SK+. Following identification of recombinants by blue/white colour selection and characterisation by restriction analysis a positive clone was identified and sequencing confirmed the integrity of the cloning sites. *NotI* restriction of the BluescriptII SK+ polylinker released the linearised 10.4kb *Ercc1* MG#13. This linearised minigene could be used directly to transfect cells. As the exact location of the promoter of the skin-specific transcript was unknown, it was possible that the 1.05kb region used in construction of the minigene did not contain all the elements required for correct expression of the skin-transcript. However, the 1.05kb region did contain the interesting (CA) and (CT) repeat regions and this was worthy of further investigation.

Figure 3.2 Construction of *Ercc1* minigene #13

a, The 5′ flanking sequence clone pRK 1.05 was cut with *SstI* and *KpnI* to liberate the insert. Drawn to scale: 1cm=100bp. b, The exon 7 to 10 region of the cDNA subclone pBluescriptII SK+ EX4-10 (RI), was isolated as a 300bp *KpnI/EcoRI* fragment. Drawn to scale 1cm=100bp. c, The 9.3kb *KpnI* genomic fragment carrying exons 1 to 7 was isolated from pKpnI 9.3. Drawn to scale: 1cm=1kb. d, Each of the isolated fragments were cloned into *SstI/EcoRI* cut pBluescriptII SK+. Drawn to scale 1cm=1kb.

Plasmid backbone shown as a thin line. 5′ flanking sequence shown in cyan, cDNA sequence shown in blue, exons shown as numbered red boxes & introns shown as thick line. Restriction sites used for cloning and subsequent analysis shown as follows: *EcoRI,* E; *BamHI,* B; *BglII*, Bg; *KpnI*, K; *PstI,* P; *SstI,* S; *NotI,* N.
Restriction sites cut for cloning or later linearising are indicated by scissors icon.
Figure adapted from original by Jim Selfridge.


SN
R
K
a.
pRK1.05
Poly(A)
E
P
K
B
E
b.
pBluescript SK+ EX4-10(RI)
K
BgB
Bg
E
P
B
Bg
E
K
1
2
3
4
5
6
7
c.
pKpnI 9.3
SN
K
BgB
Bg
E
P
B
Bg
E
K
E
1
2
3
4
5
6
7
7-10
d.
Ercc1 MG#13

3.2.2 Construction of *Ercc1* MG#11

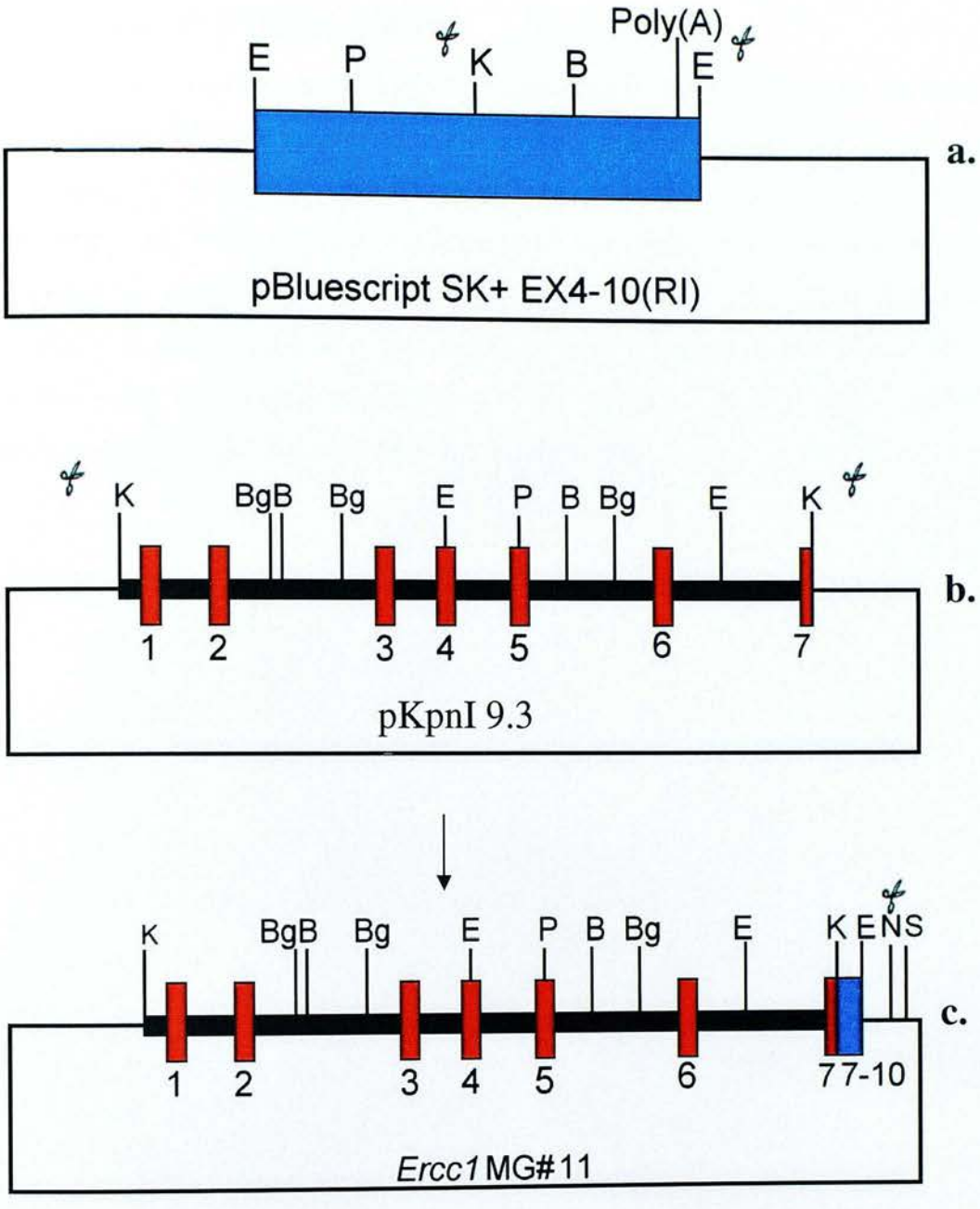
Everything upstream of a *KpnI* site in the 1.05kb flanking region was deleted in minigene#11. This deletion included the region thought to contain the skin-specific promoter and the TSS of the skin-transcript.

The following fragments were ligated to form *Ercc1* MG#11 (Figure 3.3); the fragment containing exons 1-7 (*Kpn1*) came from pKpnI9.3. This was ligated to a fragment containing exons 7-10 (*Kpn1/EcoRI*) which came from PBlueScript SK + Ex4-10(RI) in BluescriptII SK+. Ms. Carolanne McEwan (Melton Lab.) constructed *Ercc1*MG #11.

Figure 3.3 Construction of *Ercc1* minigene #11

Plasmid backbone is shown as thin line. cDNA sequence shown in blue. Exons shown as numbered red boxes. Introns shown as thick line. Restriction sites used for cloning and subsequent analysis shown as follows: *EcoRI,* E; *BamHI,* B; *BglII*, Bg; *KpnI*, K; *PstI,* P; *SstI,* S; *NotI,* N.

Restriction sites cut for cloning or later linearising are indicated by scissors icon.


Poly(A)
E
P
K
B
E
a.
pBluescript SK+ EX4-10(RI)
K
BgB
Bg
E
P
B
Bg
E
K
1
2
3
4
5
6
7
b.
pKpnI 9.3
K
BgB
Bg
E
P
B
Bg
E
K
E
NS
1
2
3
4
5
6
7
7-10
c.
Ercc1 MG#11

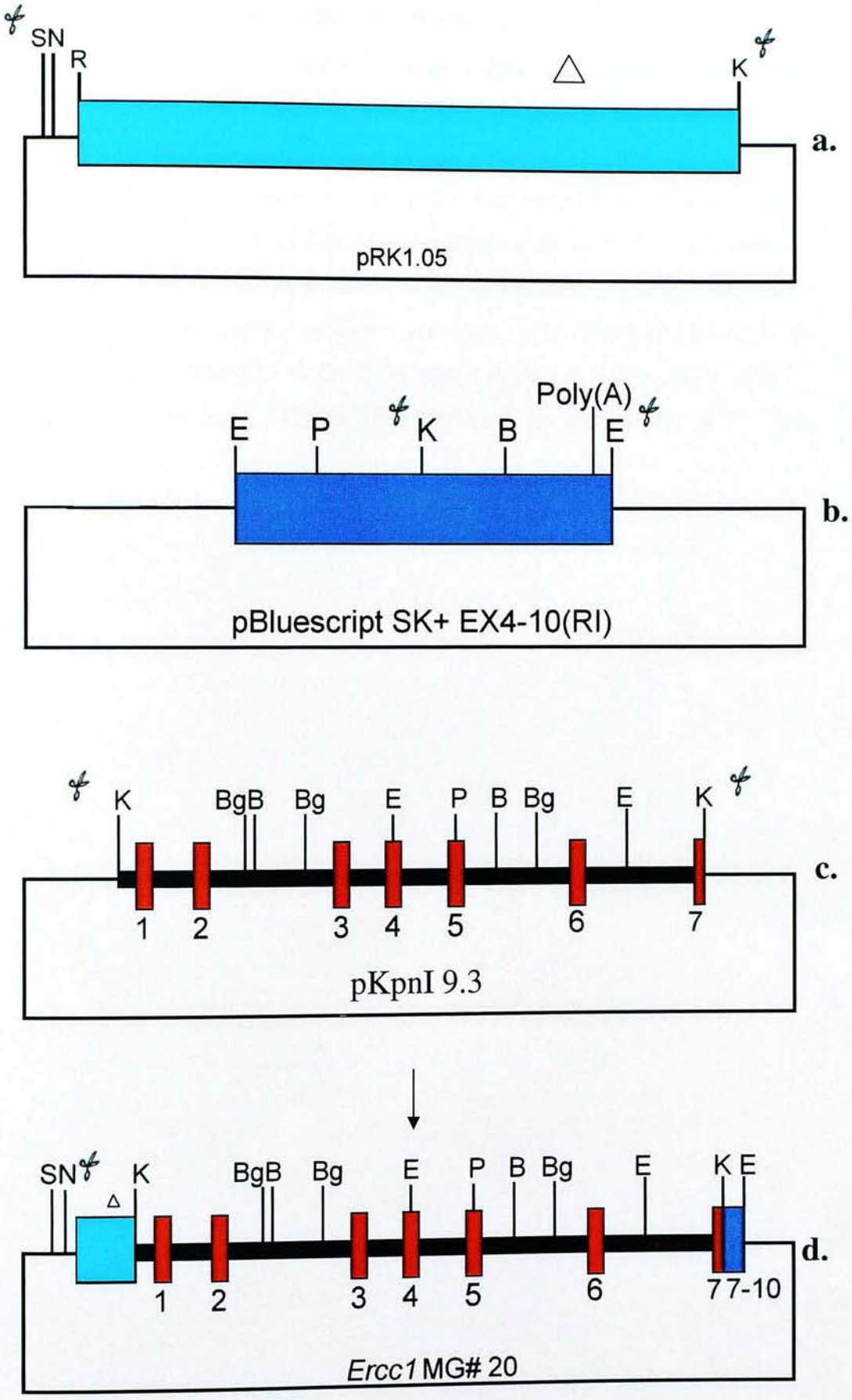
3.2.3 Construction of *Ercc1* MG#20

The following fragments were ligated in BluescriptII SK+ (*SstI/EcoRI*) to form *Ercc1* MG#20 (Figure 3.4); the 9.3kb genomic subclone pKpnI 9.3 containing exons 1-7 (*KpnI*); the 300bp fragment of the 3′ end of the *Ercc1* cDNA (*KpnI/EcoRI*) containing exons 7-10, the polyadenylation signal and a short 3′ UTR were ligated to the modified 1.05kb fragment. The (CT) repeat region of the 1.05kb fragment (*SstI/KpnI)* in BluescriptII SK+ was removed by restriction with *BseRI and BclI*. The intriguing (CT) repeat region was deleted in *Ercc1* MG#20. Ms.Carolanne McEwan (Melton Lab.) constructed *Ercc1* minigene #20.

Figure 3.4 Construction of *Ercc1* minigene #20

Plasmid backbone is shown as thin line. 5′ flanking sequence shown in cyan, cDNA sequence shown in blue. Exons shown as numbered red boxes. Introns shown as thick line. Deletion indicated by triangle. Restriction sites used for cloning and subsequent analysis shown as follows: *EcoRI,* E; *BamHI,* B; *BglII,* Bg; *KpnI,* K; *PstI,* P; *SstI,* S; *NotI,* N.

Restriction sites cut for cloning or later linearising are indicated by scissors icon.

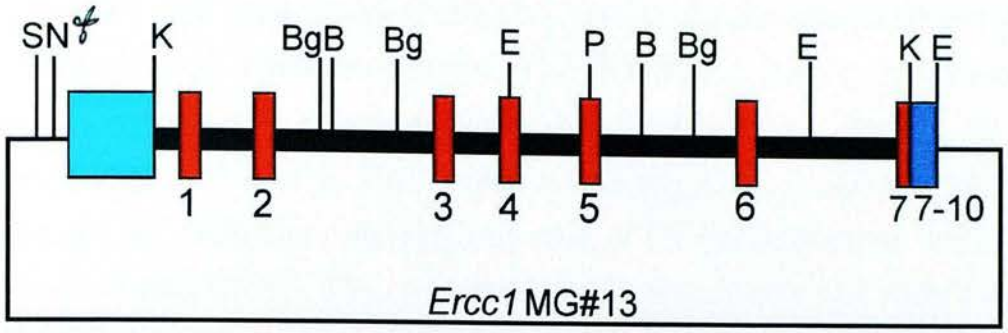

SN
R
K
a.
pRK1.05
Poly(A)
E
P
K
B
E
b.
pBluescript SK+ EX4-10(RI)
K
BgB
Bg
E
P
B
Bg
E
K
1
2
3
4
5
6
7
c.
pKpnI 9.3
SN
K
BgB
Bg
E
P
B
Bg
E
K
E
1
2
3
4
5
6
7 7-10
d.
Ercc1 MG# 20

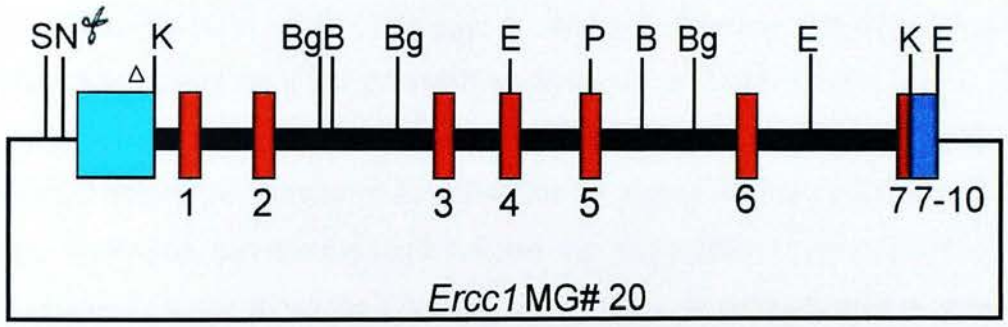
3.2.4 Summary of differences between minigenes

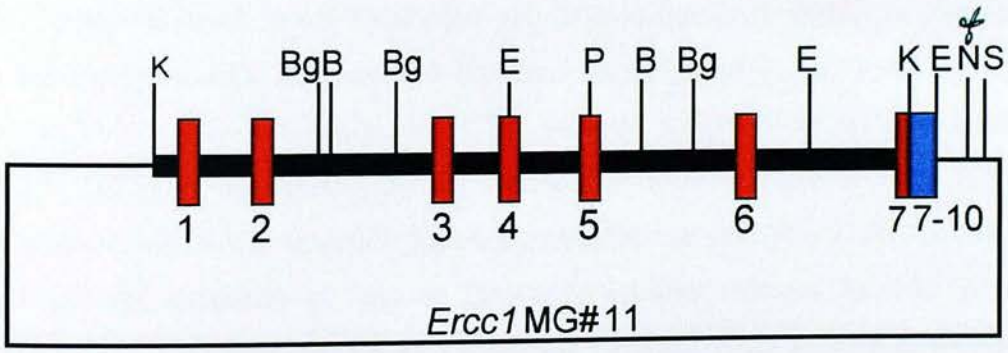
Ercc1 MG#13 contains the 1.05kb flanking region that contains the normal transcript's promoter and is likely to contain the skin-specific transcript promoter. *Ercc1* MG#11 contains the normal transcript promoter but lacks the region upstream of the *Kpn1* site that is thought to contain the skin specific transcript promoter. *Ercc1* MG#20 is very similar to *Ercc1* MG#13 apart from the (CT) repeat deletion. The (CT) repeat deletion may result in loss of skin-transcript expression if this region forms part of the skin specific transcript's promoter. Any differences between the expression of *Ercc1* MG#13 and *Ercc1* MG#20 will be due to the absence of the (CT) repeat region in *Ercc1* MG#20. See figure 3.5 for comparison of the three minigenes.

Figure 3.5 Comparison of *Ercc1* minigenes 13, 20 and 11

The three *Ercc1* minigenes, 13, 11 and 20 are shown. The plasmid backbone is shown as a thin line. 5′ flanking sequence shown in cyan, cDNA sequence shown in blue. Exons shown as numbered red boxes. Introns shown as thick line. Deletion in *Ercc1* MG#20 indicated by triangle in 5′ flanking region.


SN
K
BgB
Bg
E
P
B
Bg
E
K E
1
2
3
4
5
6
7 7-10
Ercc1 MG#13


SN
K
Δ
BgB
Bg
E
P
B
Bg
E
K E
1
2
3
4
5
6
7 7-10
Ercc1 MG# 20


K
BgB
Bg
E
P
B
Bg
E
K E
N
S
1
2
3
4
5
6
7 7-10
Ercc1 MG#11

3.3 Correction of UV hypersensitivity in murine embryonic fibroblasts transformed with Ercc1 MG#13, #20 and #11

Cells were electroporated as explained in method section. The 3:1 ratio of MG DNA to *pBT/PGK- Hprt* used for the electroporation ensured a high likelihood that *Hprt* DNA entering cells would be accompanied by MG DNA. The experimental control was electroporation of cells with *pBT/PGK- Hprt* DNA only. Subsequent co-selection was by means of HAT selection (selecting cells containing *pBT/PGK-Hprt*) and UV (selecting cells with correction of UV hypersensitivity due to the presence of MG DNA). The co-selection selects cells transformed with both the *pBT/PGK- Hprt* and *Ercc1* MG DNA. The UV-C dose chosen for post-electroporation selection was $7.5J/m^{-2}$. At this dose non-transfected cells die while cells with functional minigene survive. The UV selection dose was set to allow cells with partial function of the minigene to survive, expression merely had to be sufficient to correct the UV hypersensitive phenotype of the PF24 cells.

Table 3.2 shows the number of colonies/dish for plating density of $5X10^5$. Colony counts following transfection demonstrated that the control (*Hprt* only cells) had few, if any, colonies following UV/HAT co-selection. In contrast, cells transformed with *Hprt* DNA + *Ercc1* minigene #13, #11 or #20 survive selection and give rise to numerous colonies. The ratio of colonies in the HAT + UV dishes to colonies in the HAT dishes is much lower in the *Hprt* only transfection (ratio 0.002) in comparison to the minigene #13, #11 and #20 transfections (ratios of 0.208, 0.165 and 0.253 respectively. Clones for subsequent UV survival studies were selected from the $5X10^5$ dishes as this plating density resulted in discrete colonies. The fact that colonies survived UV selection following transfection with *Ercc1* MG#13, #11 or #20 and the similarity in ratio of HAT+UV selected colonies to HAT selected colonies for each minigene demonstrates that all three minigenes are functional in PF24 and, at this level of analysis, are capable of correcting the UV hypersensitive phenotype of the PF24 cells equally effectively.

Table 3.2 Colony counts following *Ercc1*-minigene-transfection and HAT/UV selection in *Ercc1*-deficient mouse embryonic fibroblasts

Number of colonies/petri dish for plating density of $5X10^5$. Colonies were counted following selection with HAT or co-selection with HAT + UV. UV colony numbers are an average of 3 dishes. The ratio of HAT+UV selected colonies to HAT only selected colonies is shown to indicate the ratio of cells surviving co-selection to those surviving HAT only selection.

Electroporation DNA	**No. Colonies following HAT selection (n=1)**	**No. colonies following HAT+UV selection (average of 3 dishes)**	**Ratio of HAT+UV selected colonies to HAT selected colonies**
Ercc1 MG#13 + *pBT/PGK- Hprt*	475	99 RANGE 48-197 SD ± 69.5 SE 40.2	0.208
Ercc1 MG#11 + *pBT/PGK- Hprt*	615	102 RANGE 50-84 SD ± 58.8 SE 34	0.165
Ercc1 MG#20 + *pBT/PGK- Hprt*	423	107 RANGE 61-84 SD ± 54.8 SE 31.6	0.253
pBT/PGK- Hprt	188	0.3 RANGE 0-1 SD ± 0.5 SE 0.3	0.002

Note: Without repeating the experiment with a greater number of dishes plated with cells for each transfection, it is not possible to comment upon the statistical significance of the differences in the ratio of HAT+UV to HAT selected colonies. It is possible that although the ratio of HAT + UV selected to HAT selected colonies in cells transformed with MG#11 is lower than that of the cells transformed with MG#13 and MG#20 this may not be statistically significant. The colony count for *pBT/PGK - Hprt* transfected cells following HAT selection is lower than expected when compared to the other transfections and a possible explanation is low transformation efficiency.

3.3.1 Correction of UV hypersensitivity in murine embryonic fibroblasts transformed with *Ercc1* MG#13

Before performing UV survival assays with *Ercc1* minigene#13-derived minigenes, it was necessary to examine the ability of the *Ercc1* MG#13 to correct the UV hypersensitivity of the *Ercc1* deficient murine embryonic fibroblasts, PF24. Correction of the hypersensitivity of the fibroblasts would indicate that transformation with *Ercc1* MG#13 results in the production of functional ERCC1 protein. If this minigene was not able to correct the UV hypersensitivity of the PF24 cells then continuing analysis by transforming cells with the minigenes derived from MG#13 would be pointless. Lack of functionality with the derived minigenes would provide no insights if they were produced from a non-functional minigene.

The UV survival assays were performed as follows; 16 30mm diameter petri dishes were plated with $2X10^4$ cells per dish and cultured overnight to allow the cells to settle and adhere to the surface of the dish in a uniform layer. 24 hours after plating, the dishes were irradiated with various doses of UV-C. The UV-C doses used in the survival assays cover a range which is normally not lethal to cells with wild type *Ercc1* expression while *Ercc1* null cells are killed by the higher UV doses of the survival assay. In this way, correction of UV hypersensitivity by *Ercc1* minigene expression can be ascertained by comparison of the UV survival curve of transformed cells with those of non-transformed and wild type cells.

For each UV dose, two dishes were irradiated. The experiment was performed in duplicate to provide an average survival value for each dose. In addition to the irradiated dishes, a number of control dishes were also cultured but not irradiated. All the dishes were cultured until the control dishes became confluent; at this time all the dishes were fixed and stained. The UV survival curves for the PF24 cells transformed with *Ercc1*MG#13 are shown in figure 3.6. The staining of the cells enabled quantification of survival with the stain value of the control dishes representing 100% on the survival index and percentage staining in the irradiated dishes with comparison to the controls equating to the percentage score on the survival index for each dose of UV. Cell properties that contribute to the survival

index score are the resistance of the cells to the UV dose (the percentage of cells initially surviving UV irradiation), the ability of the cells to repair damaged DNA (which leads to either apoptosis or release from cell cycle arrest), rate of DNA damage repair following irradiation and the growth rate of the surviving cells.

The plots demonstrate that the PF24 cells transfected with *Ercc1* MG#13 are not hypersensitive to UV. The survival curve of the PF24 cells which were not transfected with *Ercc1* MG#13 were hypersensitive to UV with 99.5% of cells surviving a UV dose of 0.625 Jm^{-2}, 73.4% surviving the 1.25 Jm^{-2} dose, 26.7% surviving the 2.5 Jm^{-2} dose, 9.1% surviving the 3.75 Jm^{-2} dose and 0.7% surviving the highest dose of 10 Jm^{-2}. In comparison, 84.2% of the wild type PF20 cells survived the 3.75 Jm^{-2} dose, 74.2% survived the 5 Jm^{-2} dose and 16% survived the highest dose of 10 Jm^{-2}. When the survival curves of the PF24 clones transfected with *Ercc1* MG#13 are compared to the PF20 wild type cells then the similarity is clear, indicating correction of UV hypersensitivity of the PF24 cells by transfection with *Ercc1*MG#13. MG#13 clones 1 and 2 have higher survival than the wt PF20 at the high UV doses and MG#13 clone 6's survival matches that of the wt PF20 line while MG#13 clone 8 is lower that that of PF20 and the other *Ercc1* MG#13 clones at the highest UV dose.

The survival profile of PF24 MG#13 clone 2 showed resistance to UV at both 3.75 Jm^{-2} and 5 Jm^{-2}, dropping to 48% and 25% survival at 7.5 Jm^{-2} and 10 Jm^{-2}. PF24 MG#13 clone 1 had slightly higher UV survival than clone 2 at the lower UV doses, 91.5% and 84.7% at 3.75 and 5 Jm^{-2} and slightly lower survival than clone 2 at the higher UV doses with 46% and 18% survival at 7.5 Jm^{-2} and 10 Jm^{-2}. The survival curve of PF24 MG#13 clone 6 was slightly lower than that of PF24 MG#13 clones 1 and 2, but almost identical to that of the wt PF20 with 60% and 57% survival at UV doses of 3.75 Jm^{-2} and 5 Jm^{-2} and survival of 39% and 16% at the higher UV doses of 7.5 Jm^{-2} and 10 Jm^{-2}. The survival curve of PF24 MG#13 clone 8 was lower than the other 3 clones and the PF20 cells but higher than the non-transfected cells. At 3.75 and 5 J/m^{-2}, 72 and 65% of cells survived while at 7.5 and 10Jm^{-2} 32 and 7.1% of the cells survived.

In figure 3.6, although the error bars do not quite overlap, the difference between clone 8 and the mean of clones 1,2 and 6 (mean = 24.4) is considered to be not quite statistically significant when tested using a two-tailed *t*-test (P=0.056) at UV dose $10Jm^{-2}$. When clone 8 is compared to individual clones, there is no statistically significant difference between clone 8 and clone 1 (P=0.3232). However, there is a statistically significant difference between clone 8 and clone 2 (P=0.0125) and between clone 8 and clone 6 (P=0.0154). The error bar on clone 1 is large and almost overlaps with clone 8, while the error bars on clones 8, 2 and 6 are small and do not overlap. At the lower UV doses the error bars of all the clones overlap, indicating that the difference between the clones only manifested at the highest UV dose.

From the survival curves and magnitude of the error bars on those curves it was clear that transfection of *Ercc1*-null murine fibroblasts, PF24, resulted in correction of UV hypersensitivity to a level near or equal to wt PF20 levels in all the clones. This indicates that minigene#13 transfection resulted in translation of a functional ERCC1 protein. Subsequent RT-PCR analysis (figure 3.10) enabled identification of the *Ercc1* transcripts being produced by the transfected cells.

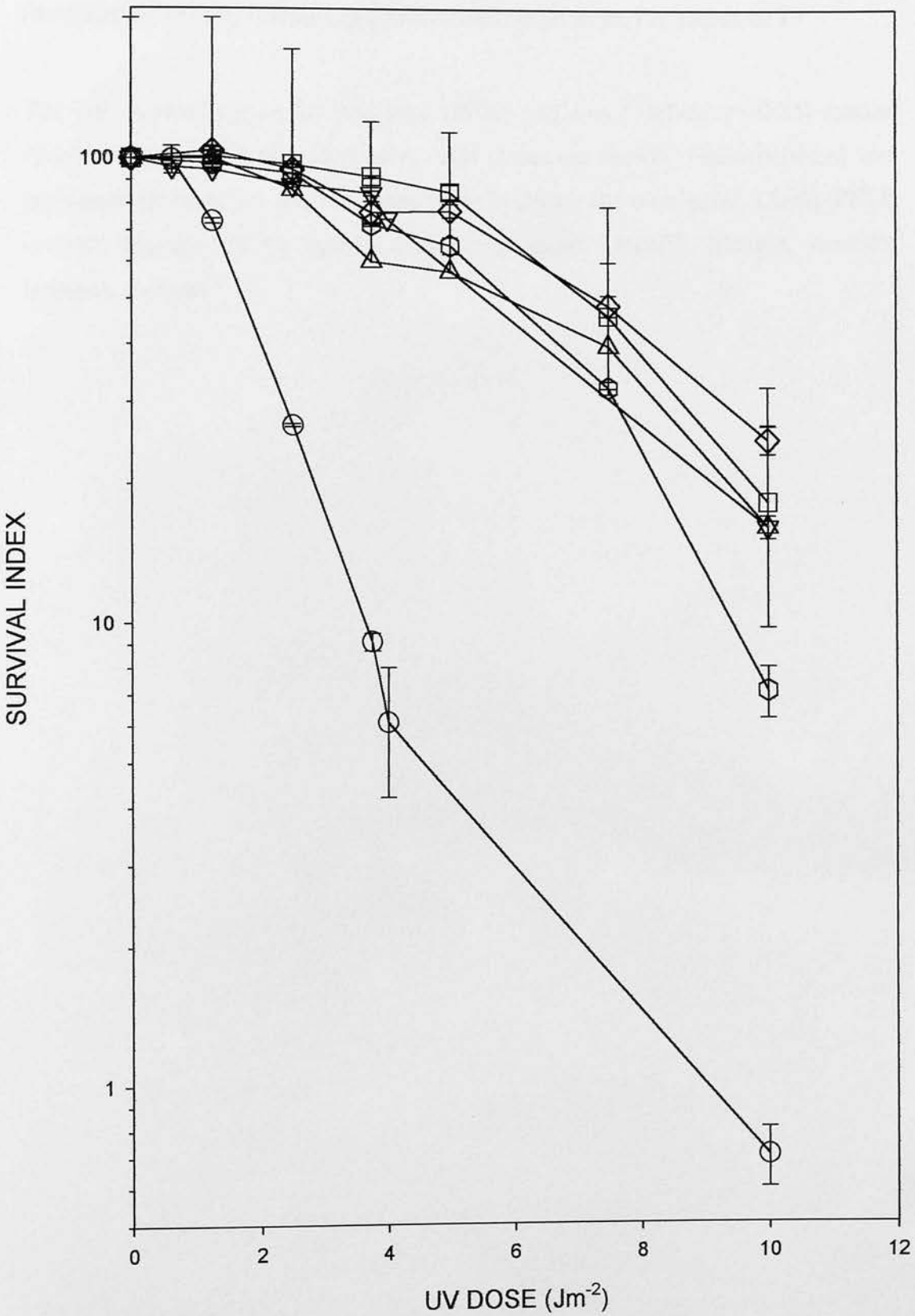

SURVIVAL INDEX
100
10
1
0
2
4
6
8
10
12
UV DOSE (Jm-2)

Figure 3.6 Correction of UV hypersensitivity of *Ercc1* deficient fibroblasts (PF24) following transfection with *Ercc1* minigene#13

The UV survival curves for wild type (PF20) and *Ercc1* deficient (PF24) murine fibroblasts and four minigene-positive PF24 clones are shown. The experiment was performed in duplicate and the mean value is shown for each point. Circle, PF24; inverted triangle, PF20; square, clone#1; diamond, clone#2; triangle, clone#6; hexagon, clone#8.

3.3.2 Expression of *Ercc1* minigene #11 in *Ercc1*-deficient murine fibroblasts, (PF24).

Ercc1 minigene#11 was produced by deletion of the region upstream of the *Kpn1* restriction site in MG#13. This means that minigene#11 does not contain the region in which the skin promoter is thought to lie, nor the transcriptional start site of the skin transcript, meaning expression of the minigene should only result in production of the normal 1.1kb transcript. Hypothetically, this could result in reduced UV survival in comparison to the cells transformed with *Ercc1* MG#13 if the skin-specific transcript contributes to UV resistance. It is also possible that deletion of this region could result in loss of functionality of the minigene. The inability to express the 1.5kb transcript due to the deletion in *Ercc1*MG#11 is more likely to have an effect upon the UV survival of transformed keratinocytes than mouse embryonic fibroblasts. This is because the fibroblasts normally produce the 1.5kb and 1.1kb transcripts at an equal ratio whereas keratinocytes produce primarily the 1.5kb transcript. This is why three different cell lines were used for investigation of the minigenes, each cell line has a different wild type *Ercc1* expression pattern and transformation with the minigenes provides different insights into the importance of the deleted regions to expression of the novel transcript.

The survival curves of the clones transfected with *Ercc1*MG#11 are shown in figure 3.7. The plot demonstrated that transfection with MG#11 corrects the UV hypersensitivity of the PF24 murine fibroblasts to wild type levels. Comparison of the mean survival of the set of clones transfected with MG#11 to the mean survival of the set of clones transfected with MG#13 indicates that the two minigenes correct UV hypersensitivity to a similar extent. The wt and null controls on the plot are the same as in the minigene#13 plot and the experiment was performed as before with only the minigene used for transfection differing.

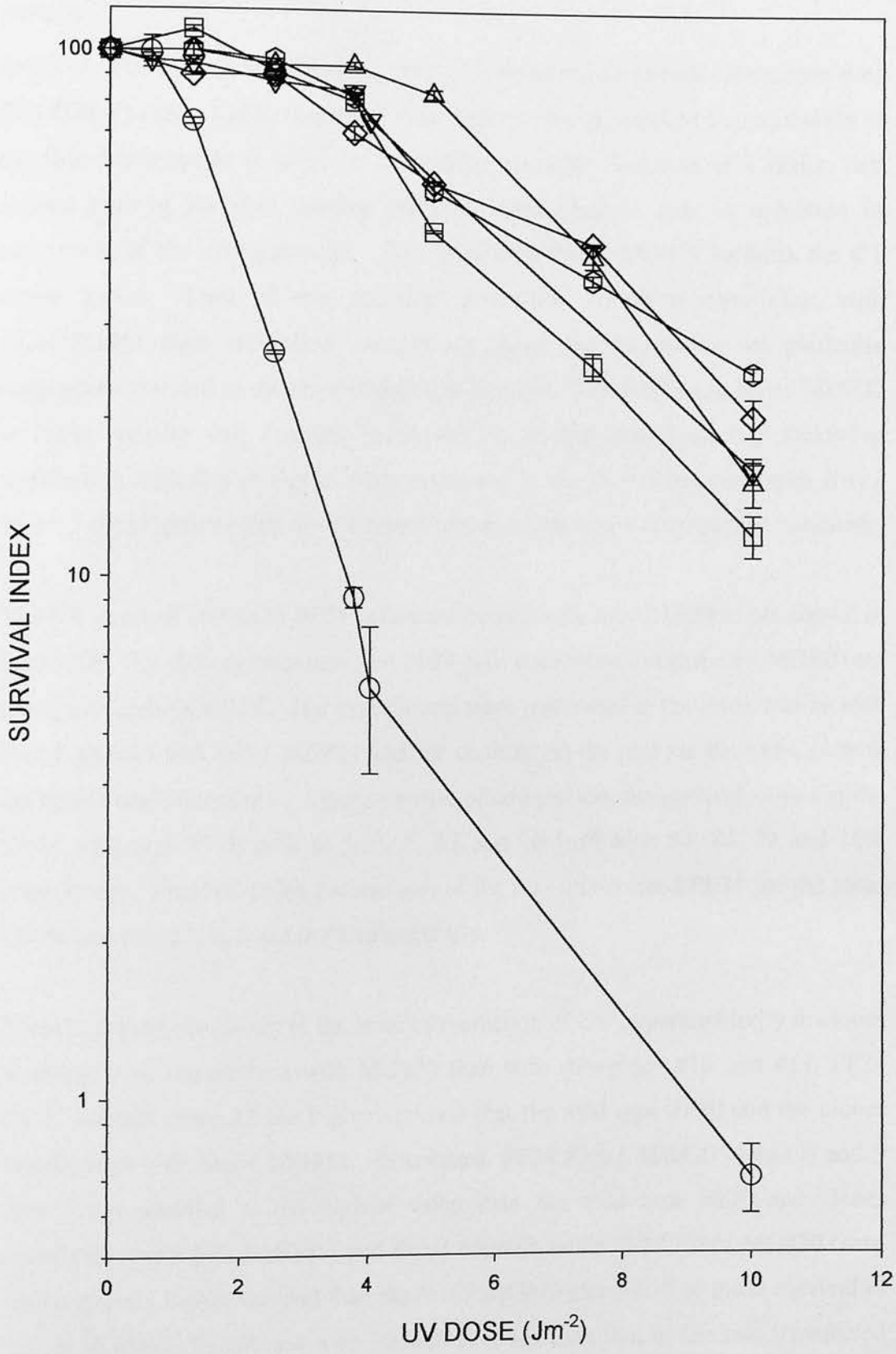
For ease of comparison, the survival values for the controls were as follows; PF20 survival at 3.75, 5, 7.5 and 10 Jm^{-2} were 88, 84, 74 and 16% respectively while survival of the non-transfected PF24 for those UV doses were 27, 9, 6 and 0.7% respectively. The survival curves of two of the clones transfected with

*Ercc1*MG#11, clones 9 and 14, were slightly higher than the wt PF20 while the survival curves of the other two clones, clones 6 and 10, were slightly lower than the wt curve. PF24 *Ercc1*MG#11 clone 14 had the highest survival curve with 82, 54, 36 and 24% surviving UV doses of 3.75, 5, 7.5 and 10 Jm^{-2} respectively. PF24 *Ercc1*MG#11 clone 9 had the second highest survival curve with 70, 56, 43 and 20% of cells surviving those UV doses. PF24 *Ercc1* MG#11 clone 10 had very slightly lower survival than the wild type PF20 with 94, 83, 41 and 15% of cells surviving UV doses of 3.75, 5, 7.5 and 10 Jm^{-2}. PF24 *Ercc1*MG#11 clone 6 had a slightly lower survival curve than clone 10 with 81, 46, 25 and 12% survival for the same UV doses.

Because the UV survival curves of the clones transformed with *Ercc1*MG#11 have a similar mean survival to the mean survival of the clones transformed with *Ercc1* MG#13, deletion of the region upstream of the *Kpn1*site (containing the skin transcript's TSS and likely promoter region) did not result in reduced ability to repair UV-induced DNA damage. RT-PCR analysis was performed to ensure that only the normal *Ercc1* transcript was being produced by the transfected cells.

Figure 3.7 Correction of UV hypersensitivity of *Ercc1* deficient fibroblasts (PF24) following transfection with *Ercc1* minigene#11

The UV survival curves for wild type (PF20) and *Ercc1* deficient (PF24) murine fibroblasts and four minigene-positive PF24 clones are shown. The experiment was performed in duplicate and the mean value is shown for each point. Circle, PF24; inverted triangle, PF20; square, clone#6; diamond, clone#9; triangle, clone#10; hexagon, clone#14.


100
10
1
SURVIVAL INDEX
0
2
4
6
8
10
12
UV DOSE (Jm-2)

3.3.3 Expression of *Ercc1* minigene #20 in *Ercc1*-deficient murine fibroblasts, (PF24).

In *Ercc1* MG#20 there is a deletion of the (CT) repeat region between restriction sites *(BseRI/BclI)* of the 1.05kb fragment. This region could be required for expression of the skin transcript as it could be a promoter element. Deletion of a region that formed part of the skin specific promoter could lead to loss or reduction in expression of the skin transcript. The deletion in *Ercc1* MG#11 includes the CT repeat region. Lack of skin transcript production following transfection with *Ercc1*MG#11 does not allow conclusions about the importance of particular sequences contained in the deleted region to be made. The deletion in *Ercc1* MG#20 is more specific and changes in expression of the skin transcript following transfection with this minigene when compared to the expression seen with *Ercc1* MG#13 could indicate that the CT repeat region is part of the skin specific promoter.

The UV survival curves of PF24 cells transformed with *Ercc1* MG#20 are shown in figure 3.8. The plots demonstrate that PF24 cells transformed with *Ercc1* MG#20 are not hypersensitive to UV. The experiments were performed in the same way as with *Ercc1* MG#13 and *Ercc1* MG#11 and the controls on the plot are the same as with the earlier transformations. Again, for ease of comparison, the survival values of the *Ercc1* wild type PF20 cells at 3.75, 5, 7.5 and 10 Jm^{-2} were 88, 84, 74 and 16% respectively. Survival index percentages of the non-transformed PF24 for the same UV doses were 27, 9, 6 and 0.7% respectively.

There is greater variability in the level of correction of UV hypersensitivity in clones resulting from transfection with MG#20 than with minigenes #13 and #11. PF24 *Ercc1* MG#20 clone 12 has higher survival that the wild type PF20 and the clones transformed with *Ercc1* MG#13. In contrast, PF24 *Ercc1* MG#20 clones 6 and 9 have lower survival at the highest value than the wild type PF20 and clones transformed with *Ercc1* MG#11 and *Ercc1* MG#13, while PF24 *Ercc1* MG#20 clone 10 has slightly higher survival than the *Ercc1* MG#13 clone 8. The mean survival of the set of clones transfected with MG#20 is lower than that of the sets transfected with MG#13 and#11.

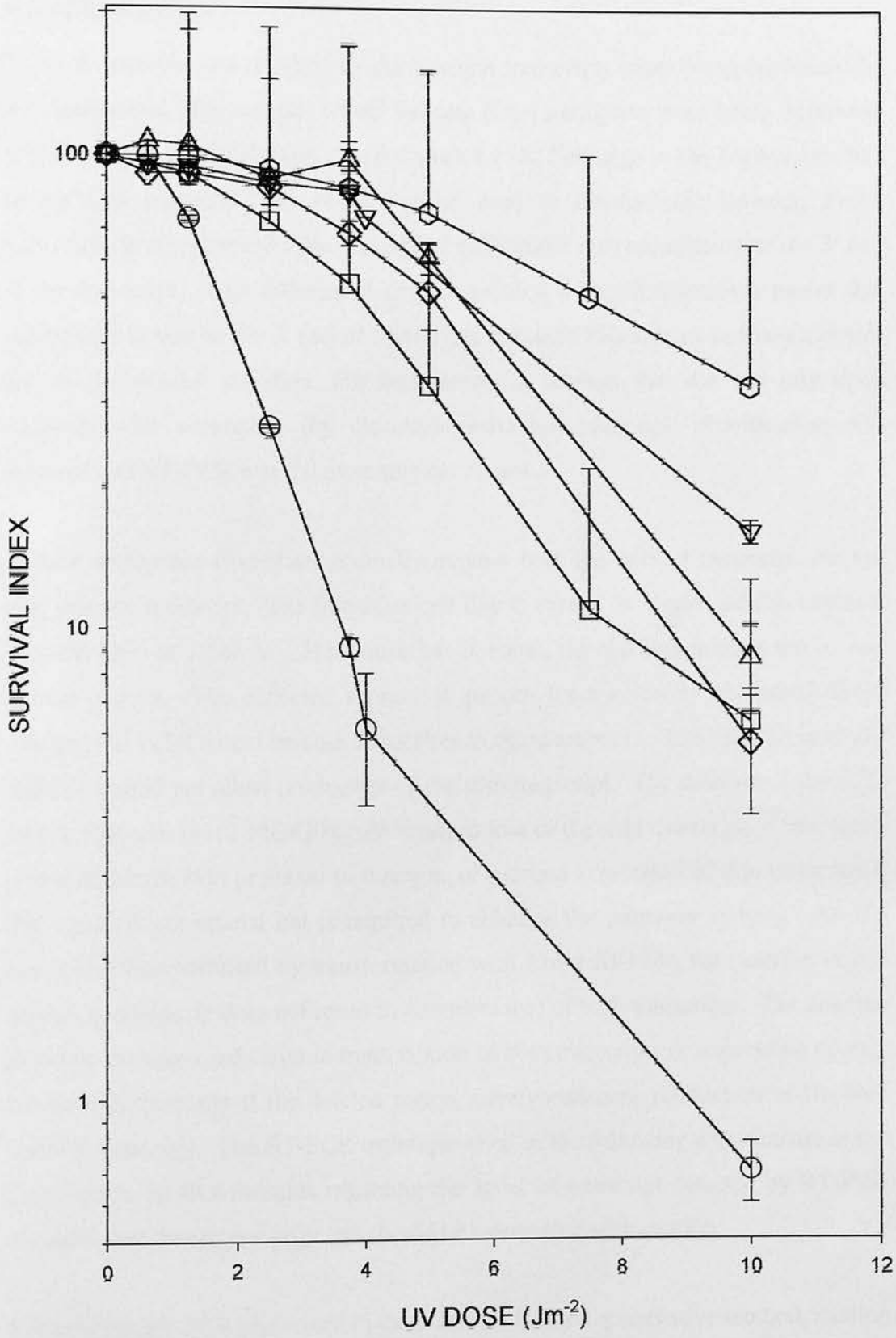
The highest survival curve was that of PF24 *Ercc1* MG#20 clone 12 with survival index percentages of 85, 76, 50 and 33% for UV doses of 3.75, 5, 7.5 and $10Jm^{-2}$ respectively. In contrast, the survival curves of PF24 *Ercc1* MG#20 clone 10 is lower than that of wild type PF20 with survival index percentages of 98, 61 and 9% survival for UV doses of 3.75, 5 and $10Jm^{-2}$ respectively. PF24 *Ercc1* MG#20 clone 9's survival curve is very low. Clone 9 survival index percentages for UV doses of 3.75, 5 and $10Jm^{-2}$ were 69, 51 and 5.75%.

Deletion of the CT repeat region in *Ercc1* MG#20 did result in reduced ability to repair UV induced DNA damage in transformed cells when compared with cells transfected with *Ercc1* MG#13 and this is indicated by the lower mean survival of the set of clones transformed with MG#20 when compared to that of the set of clones transformed with *Ercc1* MG#13.

Transfection of PF24 with minigene #13 corrects the UV hypersensitivity of PF24 to wild type (PF20) level. Loss of the region upstream of the *Kpn1* restriction site in *Ercc1*MG#11 did not impair the functionality of the minigene when compared to *Ercc1* MG#13. The discrete deletion of the CT repeat region in *Ercc1* MG#20 appears to lead to less correction of UV hypersensitivity in transformed cells when compared to *Ercc1* MG#13 and #11.

Figure 3.8 Correction of UV hypersensitivity of *Ercc1*-deficient fibroblasts (PF24) following transfection with *Ercc1* minigene#20

The UV survival curves for wild type (PF20) and *Ercc1* deficient (PF24) murine fibroblasts and four minigene-positive PF24 clones are shown. The experiment was performed in duplicate and the mean value is shown for each point. Circle, PF24; inverted triangle, PF20; square, clone#6; diamond, clone#9; triangle, clone#10; hexagon, clone#12.


100
10
1
SURVIVAL INDEX
0
2
4
6
8
10
12
UV DOSE (Jm-2)

3.4 Expression of minigenes in Ercc1 null murine fibroblasts (PF24) – RT-PCR analysis

RT-PCR analysis was required to show which transcripts were being expressed by the transformed cells and this would indicate if the minigenes were being expressed appropriately for the cell type. Earlier work by Dr. Selfridge in the Melton lab. had shown that northern blots could not be used to discriminate between *Ercc1* transcripts in transformed cells because of differential polyadenylation at the 3′ end of the transcript. This differential polyadenylation between transcripts meant that differences in size at the 3′ end of transcripts obscured differences in transcript size due to differential initiation. For this reason, a method that did not rely upon transcript size separation (by electrophoresis) for transcript identification was required and RT-PCR was the most suitable choice.

Murine embryonic fibroblasts normally express both the normal transcript and the skin specific transcript. The fibroblast cell line is dermal in origin and this explains why the ratio of 1.5kb to 1.1kb transcripts is equal, the cell line mimics the *in vivo* dermal pattern. The expected expression pattern for a correctly expressed *Ercc1* minigene in PF24 would be both transcripts in equal amounts. The deletion in *Ercc1* MG#11 should not allow production of the skin transcript. The deletion of the (CT) repeat region in *Ercc1* MG#20 could result in loss of the skin transcript, if the region is crucial for the skin promoter to function, or reduced expression of skin transcript if the region is not crucial but is required to enhance the promoter activity. As UV sensitivity was corrected by transformation with *Ercc1* MG#20, the deletion in this minigene obviously does not result in complete loss of both transcripts. The deletion could be causing a reduction in transcription of both transcripts or expression of only the normal transcript if the deleted region merely enhances production of the skin specific transcript. The RT-PCR technique used in the following experiments in not quantitative, so all comments regarding the 'level' of transcript detected by RT-PCR are presented for interest only and should be interpreted with caution.

Because the RT-PCR performed in these studies is not a quantitative method, caution must be applied when interpreting the results of transcript 'levels'. This is because

RT-PCR is highly sensitive to minute differences in initial RNA quantity because of the amplification process and due to the availability of substrate as the reaction progresses. In this thesis the expression patterns are discussed in terms of both the absence or presence of a band and in terms of the 'brightness' or level of product band for each transcript detected. Because the pattern observed between clones was reproducible, both in the case of using the same cDNA and from between batches of cDNA using the same RNA, the results regarding levels are reported for interest and should not be over-interpreted.

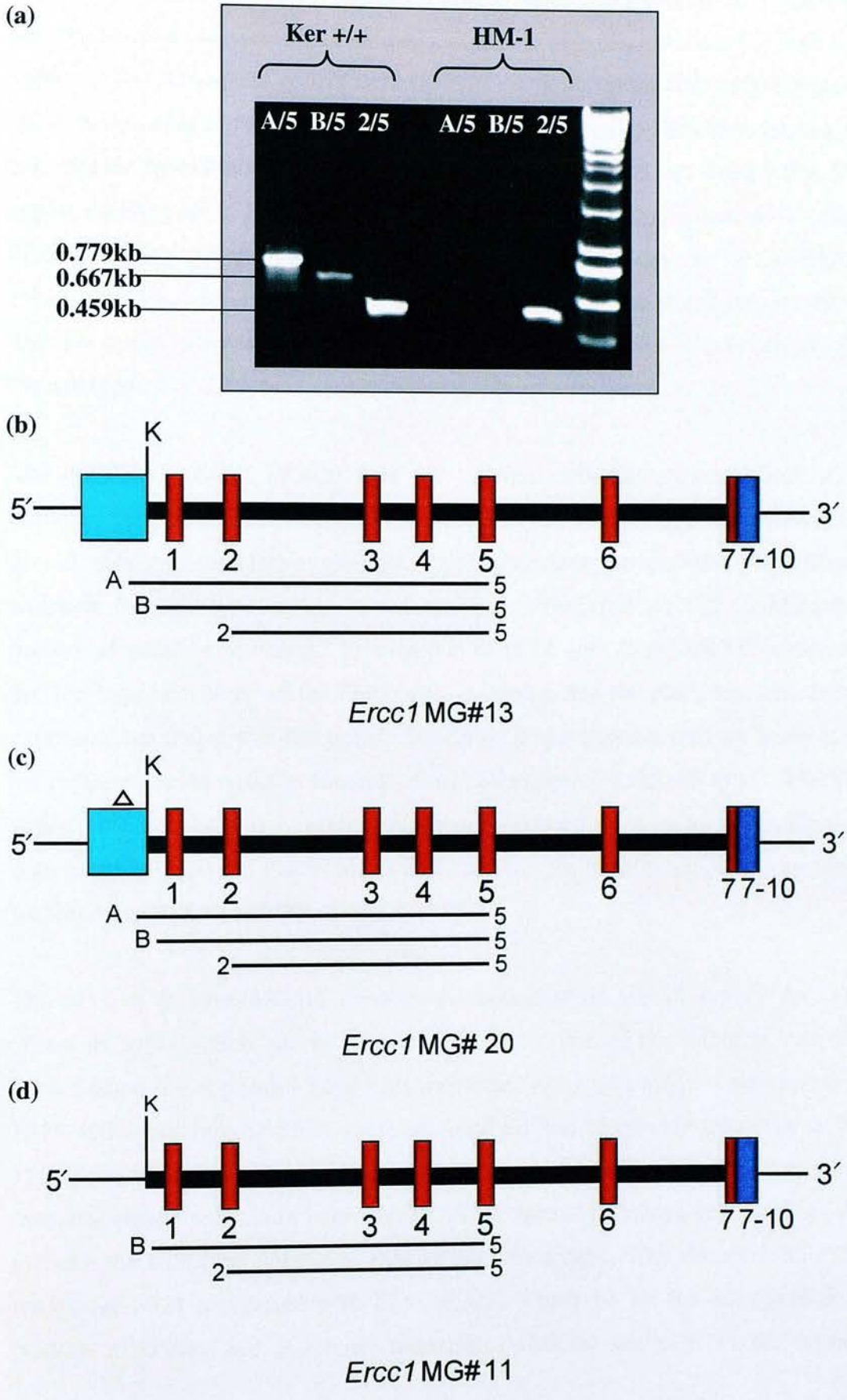
3.4.1 How *Ercc1* transcripts can be differentiated by RT-PCR

All three of the *Ercc1* minigenes were capable of correcting the UV hypersensitivity of the PF24 cells, indicating that functional *Ercc1* transcripts were being produced by the transfected cells and that the transcripts were being translated into functional protein. In order to find out whether the transformed cells were transcribing the normal *Ercc1* transcript, the skin-specific transcript or both transcripts RT-PCR was performed upon whole RNA extracted from the cells.

Four primers were used to differentiate between the two *Ercc1* transcripts. The primers with numerical prefixes are named after the number of the exon in which their hybridisation site is located. The three 5′ primers were referred to as 'A', 'B' and '2' while the 3′ primer was referred to as '5'. There are, therefore, three primer pairs A/5, B/5 and 2/5. Primer A hybridised to a region in the 1.05 flanking sequence upstream of the *Kpn1* site; Primer B hybridised to a region between the *Kpn1* site and exon 1; Primer 2 hybridised to a region in exon 2 and primer 5 hybridised to a region in exon 5. The PCR products for the primer pairs A/5, B/5 and 2/5 are 779, 667 and 459 base pairs in length respectively. The skin transcript contains the hybridisation sites of all the primers whereas the normal transcript only contains the hybridisation sites of the 2/5 primer pair. This means that when RT-PCR was performed, the skin transcript was detected by the presence of all three PCR products while the normal transcript only gave rise to the 2/5 PCR. Therefore, the A/5 & B/5 PCR products indicate the presence of the skin transcript while the 2/5 PCR product indicates the presence of normal transcript and/or skin specific transcript. A product for the 2/5 reaction without product from the A/5 reaction indicated the presence of the normal transcript in the absence of the skin specific transcript. Wild type keratinocytes produce both *Ercc1* transcripts (although almost exclusively the skin specific transcript) while embryonic cells (HM-1) produce only the normal transcript. RNA from these cells was used as the controls for the RT-PCR; in figure 3.9a all three PCR products are present in the keratinocyte sample but only the 2/5 reaction results in a product with RNA from HM-1.

Figure 3.9 Discrimination between *Ercc1* transcripts and primer hybridisation sites in *Ercc1* minigenes #11, #13 and #20

Figure 3.9 shows the three minigenes with the possible PCR products aligned to the hybridisation regions. Transfection with *Ercc1* MG#13 and #20 could hypothetically result in expression of both *Ercc1* transcripts. These would be detected by products from all three PCR reactions – see a., b. and c. *Ercc1*MG#11 does not contain the skin TSS so it is not possible for transfection with this minigene to result in expression of the skin specific transcript but production of the normal transcript is possible. This would be detected by PCR products from the 2/5 reaction while any other transcripts could be detected by the B/5 reaction – see d.


(a)
Ker +/+
HM-1
A/5 B/5 2/5
A/5 B/5 2/5
0.779kb
0.667kb
0.459kb
(b)
K
5′
3′
1 2 3 4 5 6 7 7-10
A 5
B 5
2 5
Ercc1 MG#13
(c)
K
Δ
5′
3′
1 2 3 4 5 6 7 7-10
A 5
B 5
2 5
Ercc1 MG# 20
(d)
K
5′
3′
1 2 3 4 5 6 7 7-10
B 5
2 5
Ercc1 MG# 11

Ercc1 MG#13 contains all the sequence required for hybridisation of the primers so transfection of a cell type which normally expresses both *Ercc1* transcripts with this minigene and subsequent correct expression of the minigene should, hypothetically, result in detection of two transcripts with RT-PCR (see figure 3.9b for minigene 13 and primer hybridisation). Transfection with *Ercc1* MG#13 corrected the UV hypersensitivity of the PF24, indicating that transcription and translation were taking place following transfection and that the minigene was giving rise to functional ERCC1 protein. RT-PCR would enable identification of the transcripts produced and this would indicate whether the minigene was being expressed appropriately for the cell type.

The deletion in *Ercc1* MG#20 does not interfere with the primer hybridisation sequence but it could affect transcript production if the deleted sequence is necessary for skin transcript production (see figure 3.9c). Therefore, transfection of fibroblasts, which in the wild type express both transcripts, with *Ercc1* MG#20 could have a number of possible outcomes. Transfection of PF24 with *Ercc1* MG#20 corrected the UV hypersensitivity of the fibroblasts, indicating that the minigene was being expressed and that it was functional. However, it was possible that the deletion in the minigene could result in abnormal *Ercc1* expression for the cell type. RT-PCR would indicate which transcript(s) were being produced in response to transfection with *Ercc1* MG#20 and this would indicate whether the deleted region was required for skin transcript production.

The deletion in *Ercc1*MG#11 removes the hybridisation site of primer A. The primer B hybridisation site is very close to the 5′ end of the minigene and the hybridisation site of primers 2 and 5 are unaffected (see figure 3.9d). Transfection of PF24 with this minigene could not result in production of the skin transcript as the TSS of this transcript was deleted from the minigene. Transfection of PF24 with this minigene should only result in production of the normal transcript, regardless of the fact that the wild type fibroblasts express both transcripts. The expected RT-PCR results for PF24 transfected with *Ercc1*MG#11 would be for the 2/5 reaction to produce a product and any other transcripts initiating upstream of the normal

transcript would be detected by the B/5 reaction. It would not be possible to get a product from the A/5 reaction.

See figure 3.9 (a). In wild type fibroblasts both the normal and the skin specific transcripts are produced at equal levels. In contrast, wild type keratinocytes express the skin specific transcript at a much higher level than fibroblasts do. When RT-PCR is performed on these cell types the resulting A/5 and 2/5 product bands in fibroblasts are approximately equal in intensity while in keratinocytes the A/5 band is brighter than the 2/5 band. The B/5 reaction is not as efficient as the A/5 and 2/5 reactions so this band is less intense than the A/5 band although both reactions detect the skin specific transcript. Only the 2/5 reaction results in a band when RNA from embryonic HM-1 cells is tested and this is because these cells only express the normal transcript. Untransformed PF24 cells do not result in any signal when RT-PCR is performed. Keratinocytes and HM-1 cells act as controls for the RT-PCR reactions.

3.4.2 Expression of *Ercc1* MG#13 in *Ercc1*-deficient murine fibroblasts (PF24) – RT-PCR analysis

The hypothetical result for PF24 fibroblasts transformed with *Ercc1* MG#13 would be that both transcripts would be produced if the minigene contained all the sequences necessary for their expression.

When the RT-PCR results were examined (figure 3.10) the expression of the minigene was appropriate for the cell type in 3 of the four clones. Clones #2, #6 and #8 each had positive signals in all three PCR reactions, indicating production of both transcripts. The keratinocyte and embryonic cell controls gave the correct expression pattern. One of the clones, clone #1, had an abnormal expression pattern as there was only one positive signal in the 2/5 reaction and this was extremely weak (it can only be seen in the higher exposure panel below the main figure). Although the experiments were repeated to ensure that the RNA levels were equal in all cases the low band intensity from this clone proved to be reproducible. Surprisingly the low expression in this clone did not correlate with lower UV survival so the *Ercc1* was

expressed at a level capable of repairing the DNA damage following UV irradiation. The three remaining clones, #2, 6 and 8 all expressed both transcripts and RT-PCR resulted in strong bands for all three reactions. The bands in clone #2 were strong in all three reactions. Clones #6 and 8 had strong A/5 and 2/5 bands while the B/5 band was a less intense. In later experiments the B/5 reaction conditions were changed from 30 cycles to 32 cycles.

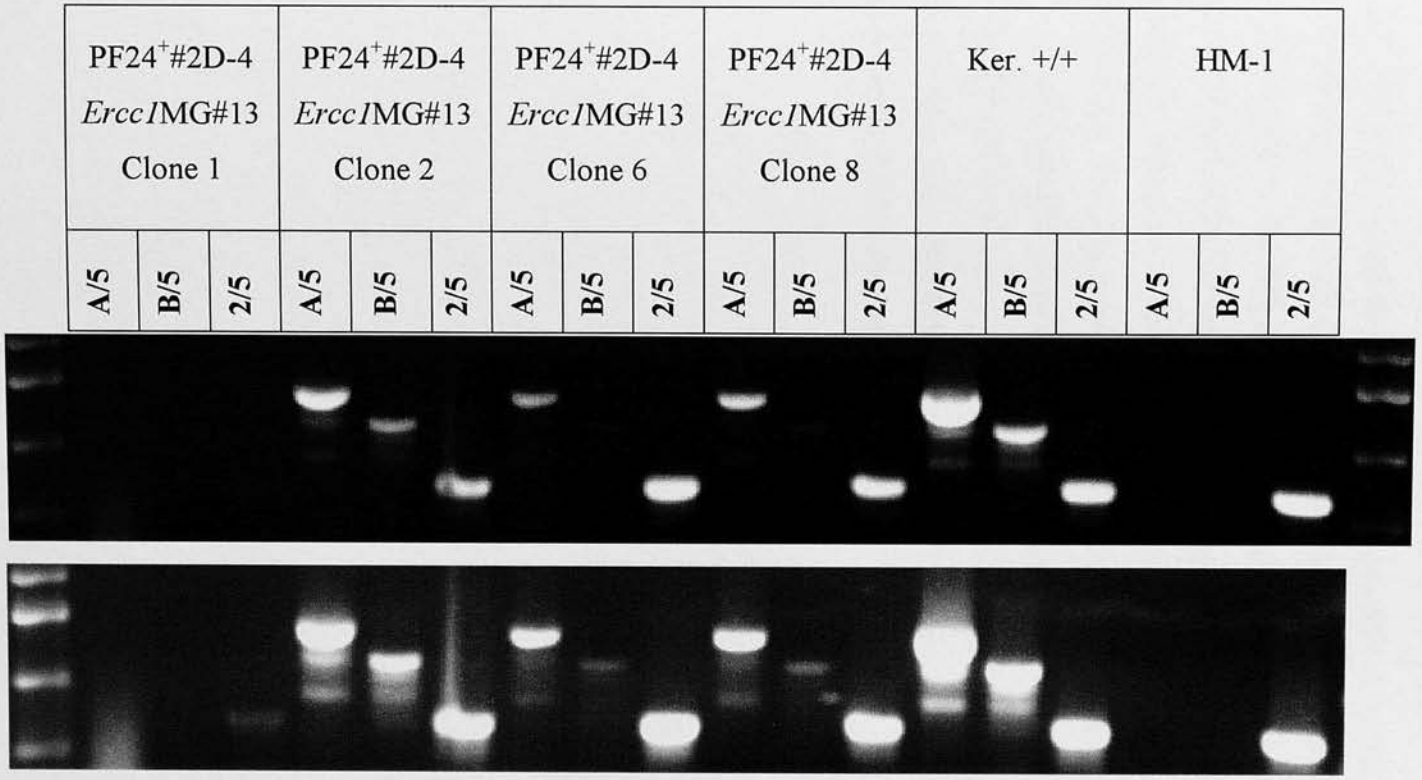
RT-PCR is not considered to be a quantitative technique so caution must be applied when interpreting any difference in the strength of the signal resulting from the PCR reactions. However, as every effort was made to ensure that the conditions were equal in all the tests, it is worth investigating whether there was any correlation between signal intensity (hypothetically representing level of *Ercc1* transcript) and UV survival. If intensity of the 2/5 band directly indicates total transcript level in the clone, this would mean that the transcript level is highest in clone 2 followed by clones 8, 6 and 1. The clones with the highest percentage UV survival at the highest UV dose (figure 3.6) are clone 2 followed by clone 1, 6 and 8. Therefore, there was no correlation between the actual intensity of the bands and the UV survival of the clones. All four clones expressed the transcripts in sufficient quantity to correct the UV hypersensitivity of the fibroblasts.

In summary, 3 clones were producing both transcripts and in each clone the transcripts were produced at roughly equal levels while one clone had an abnormal expression pattern. The expression pattern of 3 clones is correct for the cell type and indicates that *Ercc1* MG#13 is functional and contains all the elements necessary for expression of both transcripts.

Figure 3.10 Expression of *Ercc1* MG#13 in PF24 - RT-PCR

RT-PCR analysis of RNA extracted from *Ercc1*-null PF24 clones transformed with *Ercc1*MG#13. Total RNA was prepared and reverse transcribed using oligo-dT primers. The *Ercc1* cDNA was then PCR amplified from the cDNA pool of each clone. Three PCR reactions were performed on cDNA from each clone using the following forward primers respectively; primer 'A' is specific a region upstream of the normal transcriptional start site, primer 'B' is specific to a region upstream of the *Kpn1* restriction site and primer '2' is specific to exon 2. Primer '5' is specific to a region in exon 5 and is the reverse primer for each of the three reactions. 20μg of cDNA was used in each PCR reaction. The primers used were Sigma primers. The PCR was set to 30 cycles for all three PCR reactions. The PCR products were electrophoresed at 100V on an ethidium bromide/1% agarose gel for 1.5 hours.

The lower panel is a higher exposure of the gel shown in the upper panel to demonstrate the presence of a PCR product from the 2/5 reaction with clone#1.


PF24+#2D-4
Ercc1MG#13
Clone 1
PF24+#2D-4
Ercc1MG#13
Clone 2
PF24+#2D-4
Ercc1MG#13
Clone 6
PF24+#2D-4
Ercc1MG#13
Clone 8
Ker. +/+
HM-1
A/5
B/5
2/5
A/5
B/5
2/5
A/5
B/5
2/5
A/5
B/5
2/5
A/5
B/5
2/5
A/5
B/5
2/5

3.4.3 Expression of *Ercc1* MG#11 in *Ercc1*-deficient murine fibroblasts (PF24) – RT-PCR analysis

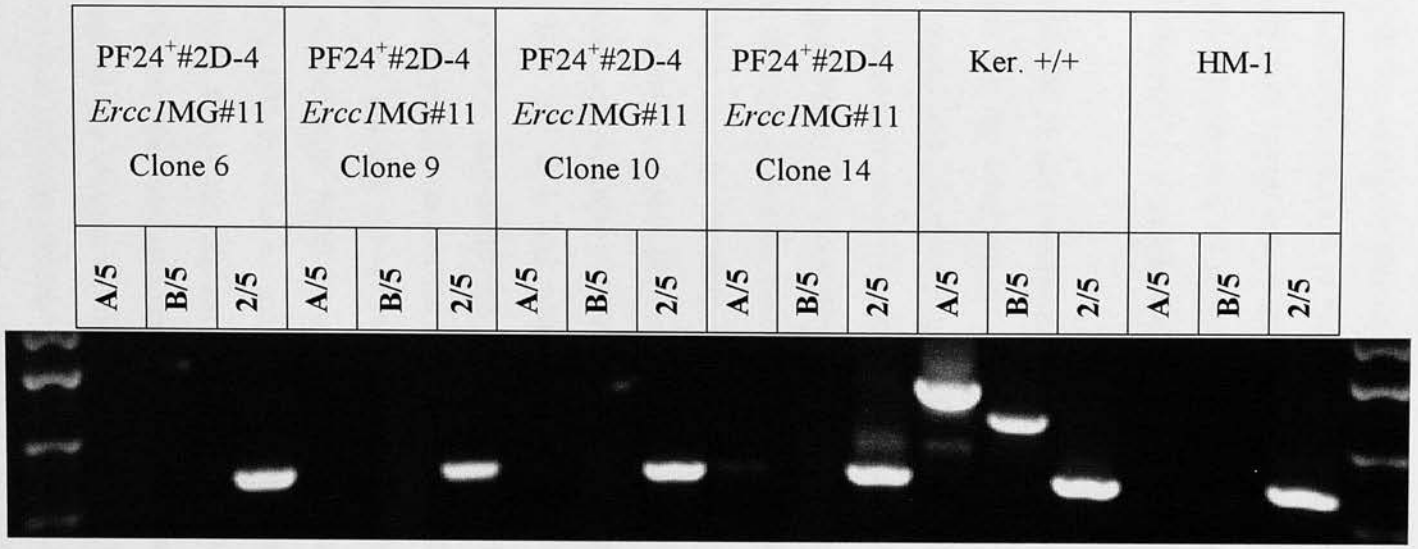
The hypothetical result for PF24 fibroblasts transformed with *Ercc1* MG#11 would be that only the normal transcript would be expressed. The RT-PCR results for fibroblasts transformed with *Ercc1*MG#11 are shown in figure 3.11. All four of the clones only had positive signals in the 2/5 reactions, indicating that only the normal transcript was expressed in the fibroblasts transformed with *Ercc1*MG#11. There is no PCR product from the B/5 reaction, indicating no aberrant upstream transcripts are being produced (this relevance of this observation will become apparent when these results are compared to keratinocytes transformed with this minigene). When it is considered that fibroblasts normally express both transcripts it is surprising that lack of the skin specific transcript did not reduce the viability of the transformed cells in the UV survival assay (figure 3.7). The UV hypersensitivity of the clones transformed with *Ercc1*MG#11 was just as effectively corrected as the clones containing *Ercc1* MG#13 (figures 3.6 and 3.7).

Figure 3.11 Expression of *Ercc1* MG#11 in PF24 - RT-PCR

RT-PCR analysis of RNA extracted from *Ercc1*-null PF24 clones transformed with *Ercc1*MG#11. Total RNA was prepared and reverse transcribed using oligo-dT primers. The *Ercc1* cDNA was then PCR amplified from the cDNA pool of each clone. Three PCR reactions were performed on cDNA from each clone using the following forward primers respectively; primer 'A' is specific to a region upstream of the normal transcriptional start site, primer 'B' is specific to a region upstream of the *Kpn1* restriction site and primer '2' is specific to exon 2. Primer '5' is specific to a region in exon 5 and is the reverse primer for each of the three reactions.

20μg of cDNA was used in each PCR reaction. The primers used were Sigma primers. The PCR was set to 30 cycles for all PCRs. The PCR products were electrophoresed at 100V on an ethidium bromide/1% agarose gel for 1.5 hours.

The PCR product in the A/5 lane of clone#14 is likely to be an over-spill from the 2/5 lane of clone#10


PF24+#2D-4 Ercc1MG#11 Clone 6
PF24+#2D-4 Ercc1MG#11 Clone 9
PF24+#2D-4 Ercc1MG#11 Clone 10
PF24+#2D-4 Ercc1MG#11 Clone 14
Ker. +/+
HM-1
A/5
B/5
2/5

3.4.4 Expression of *Ercc1* MG#20 in *Ercc1*-deficient murine fibroblasts (PF24) – RT-PCR analysis

The hypothetical result for PF24 fibroblasts transformed with *Ercc1* MG#20 depends upon the importance of the deleted (CT) repeat region to skin transcript expression. Assuming that this region is required for correct expression of the skin transcript, transformation with this minigene that lacks the region could result in no expression of the skin transcript but expression of the normal transcript or reduced expression of the skin transcripts with expression of the normal transcript. If this region were not required for correct expression of the skin transcript then deletion of the region would not impair the ability of transformed cells to express both transcripts.

RT-PCR results for the clones transformed with *Ercc1* MG#20 are shown in figure 3.12. The results for the keratinocyte and embryonic cells demonstrate that the primers were all functioning correctly in the RT-PCR. Transfection with MG#20 resulted in mixed expression of the skin specific transcript between clones, indicating that the deleted region may be important for correct expression of the skin specific transcript. Three clones had product bands for all three reactions while one only had a 2/5 signal. In terms of band intensity, only clone#6 had strong bands for all three PCR reactions. Two of the clones (#10 and 12) had weak A/5 bands with strong 2/5 bands while clone#9 only resulted in a positive signal in the 2/5 reaction. In three of the four clones both the skin and normal transcripts were being expressed, as indicated by the positive signals for the A/5 and 2/5 reactions. One of the clones, clone#9, was only expressing the normal transcript and this pattern was abnormal for the cell type. The B/5 reaction was quite variable between clones. The lower levels of A/5 reaction product were intriguing, yet when the actual intensity of the bands was correlated with UV survival no pattern emerged. In order of decreasing A/5 band intensity, clone 6 had the highest band intensity and this was followed by clones 12 and 10 (equal intensity) and finally clone 9. When the clones are arranged in decreasing order for the highest survival at the highest UV dose (figure 3.7) then the highest UV survival index percentages were in clone 12 followed by clones 10, 6 and 9. Although clone 9 did have the lowest band intensity for A/5 and the lowest UV survival the UV survival of the other clones did not correlate with the intensity

of their A/5 bands. As the RT-PCR used was not quantitative, the level of product band cannot be equated to the level of transcript.

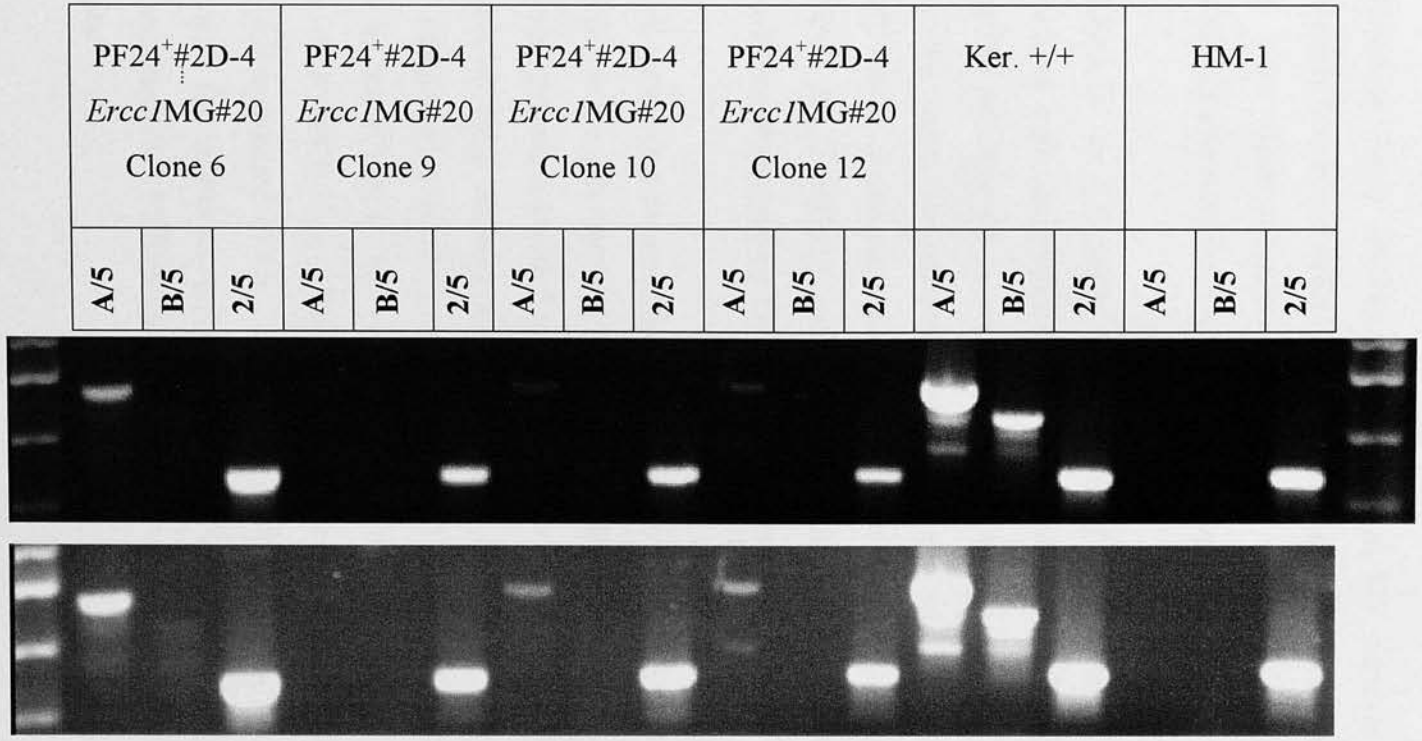
Expression of *Ercc1* MG#20 resulted in no skin specific transcript production in one of the clones but this did not result in reduced repair of UV hypersensitivity in transformed cells. The skin-specific transcript was present in the remaining three transformed clones and although the product bands were weak, the UV hypersensitivity of these clones was corrected. It appeared that removal of the (CT) repeat region may result in reduced expression of the skin transcript, especially in clone #9. However, the presence of the (CT) repeat region could not be an absolute requirement for skin transcript production as transformed clones #6, 10 and 12 still gave positive signals for the A/5 reactions.

Figure 3.12 Expression of *Ercc1* MG#20 in PF24 - RT-PCR

RT-PCR analysis of RNA extracted from *Ercc1*-null PF24 clones transformed with *Ercc1*MG#20. Total RNA was prepared and reverse transcribed using oligo-dT primers. The *Ercc1* cDNA was then PCR amplified from the cDNA pool of each clone. Three PCR reactions were performed on cDNA from each clone using the following forward primers respectively; primer 'A' is specific to a region upstream of the normal transcriptional start site, primer 'B' is specific to a region upstream of the *Kpn1* restriction site and primer '2' is specific to exon 2. Primer '5' is specific to a region in exon 5 and is the reverse primer for each of the three reactions.

20μg of cDNA was used in each PCR reaction. The primers used were Sigma primers. The PCR was set to 30 cycles for all PCR reactions. The PCR products were electrophoresed at 100V on an ethidium bromide/1% agarose gel for 1.5 hours.

The lower panel is a higher exposure of the gel shown in the upper panel to more clearly show the presence of a PCR product in the A/5 reactions of clones 10 and 12.


PF24+#2D-4
Ercc1MG#20
Clone 6
PF24+#2D-4
Ercc1MG#20
Clone 9
PF24+#2D-4
Ercc1MG#20
Clone 10
PF24+#2D-4
Ercc1MG#20
Clone 12
Ker. +/+
HM-1
A/5
B/5
2/5
A/5
B/5
2/5
A/5
B/5
2/5
A/5
B/5
2/5
A/5
B/5
2/5
A/5
B/5
2/5

3.5 Conclusions

The three *Ercc1* minigenes are functional in PF24. This is demonstrated by the survival of colonies following UV selection post electroporation, as well as by the results of the UV survival assays.

The deletion in *Ercc1* MG#11 did not reduced UV survival in comparison to transformation with *Ercc1* MG#13. Similar numbers of colonies survived UV + HAT selection post electroporation for the two minigenes. Both corrected UV hypersensitivity to a similar extent in the UV survival assays.

The deletion of the CT repeat region in *Ercc1* MG#20 reduced UV survival in comparison to cells transformed with *Ercc1* MG#13. The number of colonies surviving UV selection post electroporation was similar between the two minigenes. However, the scores for percentage survival index of UV survival curves at the higher UV doses of the survival assay was lower in cells transfected with *Ercc1* MG#20 than *Ercc1* MG#13, indicating less effective correction of the PF24 UV hypersensitivity phenotype by *Ercc1* MG#20 expression.

The *Ercc1* expression pattern observed in PF24 transformed with the minigenes was appropriate for the cell type in 3 of the 4 clones transformed with *Ercc1* MG#13 and 3 of 4 clones transfected with *Ercc1* MG#20. This result suggested that the CT repeat region deleted from MG#20 may be required for correct expression of the skin specific transcript as band intensity for the skin-specific transcript product was reduced in the RT-PCR but no conclusions may be drawn in the absence of quantitative RT-PCR data. The deletion in *Ercc1*MG#11 meant that it was not possible for expression of the 1.5kb transcript to take place so, taking the deletion into account, the expression pattern of the 4 clones transfected with *Ercc1* MG#11 was as expected (2/5 PCR product only) with no aberrant expression of upstream transcripts.

Chapter 4

Use of *Ercc1* minigenes to correct UV hypersensitivity of *Ercc1*-deficient Chinese hamster ovary cells

4.1 Expression of minigenes in cells which express only the normal Ercc1 transcript

The main purpose of transfecting cells with the *Ercc1* minigenes was the identification of sequences responsible for production of the skin transcript. For those experiments cells which would normally express both the normal 1.1kb and the skin specific 1.5kb transcripts were utilised. However, it was also important to determine if the minigenes would be expressed appropriately in other cell lines, such as those that did not normally express the skin transcript. Northern blot analysis by Dr. Selfridge (Melton lab.) had identified two cell types that expressed only the normal 1.1kb transcript; murine ES cells and Chinese hamster ovary (CHO) cells. (Figure 4.1) Since the minigenes were constructed using a murine sequence, an *Ercc1*-null murine ES cell line would have been useful for these tests. However, no such cell line was available in the Melton lab.

The human *ERCC1* gene was the first mammalian DNA repair gene to be discovered. This was by virtue of its ability to complement the UV hypersensitivity of the CHO cell line CHO43.3B. Using such a cell line would enable to cell type-specificity of the minigene expression patterns to be examined. Results would indicate whether the minigenes were expressed appropriately in a cell type that does not produce the skin specific transcript. Furthermore, the results would demonstrate the ability of the minigenes to be expressed in cells from a different species and correct the UV sensitivity of the non-murine cells.

Figure 4.1 Northern blot analysis of *Ercc1* transcript present in cultured cell lines.

Northern blot analysis of RNA samples from a range of cultured cell lines (performed by Dr. Selfridge, Melton Lab.). Total RNA was prepared from a mouse embryonic stem cell line (HM-1), a wild-type mouse fibroblast cell line (PF20), an *Ercc1* null fibroblast cell line (PF24), a wild-type Chinese hamster ovary cell line (CHO9), an *Ercc1* null Chinese hamster ovary cell line (CHO43.3B), a wild-type mouse keratinocyte cell line (Ker. (+/+)) and an *Ercc1* null mouse keratinocyte cell line, Ker. (-/-).

RNA was electrophoresed through a 1.4% agarose-formaldehyde gel, transferred onto a nylon membrane and probed with an 800bp *BamHI* fragment of the *Ercc1* cDNA, corresponding to exons 1-8. Positions of the 18S and 28S rRNAs are indicated.

HM-1
CHO9
CHO43.3B
PF20
PF24
Keratinocyte +/+
Keratinocyte -/-
1.5kb
1.1kb
28S
18S

4.2 Basis of Ercc1-deficiency in CHO43.3B (TG#1)

Fortuitously, a spontaneously occurring CHO43.3B derived cell line was available, CHO43.3B (TG#1). This cell line was both *Ercc1* and *Hprt* deficient, enabling the same selection regime to be adopted following minigene transfection as in the fibroblast experiments. CHO43.3B cells have a non-conservative point mutation that results in an amino acid substitution at the 98th Ercc1 residue. The cells are deficient in Ercc1 protein but do produce normal levels of the 1.1kb transcript (see figure 4.1) (Hayashi *et al*, 1998). This transcript was detectable on northern blots. The *Ercc1* transcript produced by the *Ercc1* knockout cells has aberrant splicing because a neo cassette disrupts exon 5. The RT-PCR assay used to detect *Ercc1* transcripts was designed specifically to distinguish between transcripts produced from minigene expression/normal expression and the endogenous transcript found in the *Ercc1* knockout CHO43.3B cells. The transcript in the CHO43.3B cells does not contain the sequence homology to primer 5 of the RT-PCR reaction.

4.3 Correction of UV hypersensitivity of Ercc1-deficient CHO43.3B (TG#1) transformed with Ercc1 MG#13, 20 and 11

As the wild type CHO cell line, CHO9, normally expresses the 1.1kb *Ercc1* transcript and very low levels of the 1.5kb transcript, transfection of CHO43.3B (TG#1) with *Ercc1* MG#13 should primarily result in production of the 1.1kb transcript if the minigene is expressed appropriately for the cell type.

Colony counts following transfection would demonstrate that the control (*Hprt* only cells) would have few, if any, colonies following UV/HAT co-selection whereas cells transformed with MG and *Hprt* DNA would survive and give rise to numerous colonies. Clones for subsequent UV survival studies were selected from the $1X10^5$ dishes as this plating density resulted in discrete colonies.

Table 4.1 shows the number of colonies/dish for a plating density of $1X10^5$. Colonies were counted following selection with HAT or co-selection with HAT and UV. The ratio of HAT+UV selected colonies to HAT only selected colonies is shown

to indicate the ratio of cells surviving co-selection to those surviving HAT only selection. The ratio of HAT+UV selected colonies to HAT selected colonies was similar between the three minigenes, as were the colony numbers.

Table 4.1 Colony counts following *Ercc1*-minigene-transfection and HAT/UV selection in *Ercc1*-deficient CHO43.3B (TG#1) cells

Number of colonies/petri dish for plating density of $1X10^5$. Colonies were counted following selection with HAT or co-selection with HAT + UV. UV colony numbers are an average of 2 dishes. The ratio of HAT + UV selected colonies to HAT only selected colonies is shown to indicate the ratio of cells surviving co-selection to those surviving HAT only selection

Electroporation DNA	No. Colonies following HAT selection	No. colonies following HAT+UV selection	Ratio of HAT+UV selected colonies to HAT selected colonies
Ercc1 MG#13 + *pBT/PGK-Hprt*	30	13 (n=3) Range 7-24 SD ± 7.6 SE 4.4	0.433
Ercc1 MG#11 + *pBT/PGK-Hprt*	57	23 (n=3) Range 16-27 SD ± 4.8 SE 2.8	0.404
Ercc1 MG#20 + *pBT/PGK-Hprt*	36	22 (n=3) Range 16-27 SD ± 4.6 SE 2.7	0.611
pBT/PGK-Hprt	59 (n=2) Range 54-64 SD ± 5 SE 3.5	0.5 (n=2) Range 0-1 SD ± 0.5 SE 0.4	0.009

Note: Without repeating the experiment with a greater number of dishes plated with cells for each transfection, it is not possible to comment upon the statistical significance of the differences in the ratio of HAT+UV to HAT selected colonies. It is possible that although the ratio of HAT + UV selected to HAT selected colonies in cells transformed with MG#11 and MG#20 is lower than that of the cells transformed with MG#13 this may not be statistically significant.

4.3.1 Correction of UV hypersensitivity in CHO43.3B (TG#1) transformed with *Ercc1* MG#13

As in the fibroblast experiments, cells were co-transfected with a 3:1 ratio of *NotI* linearised *Ercc1 MG#13* and the PGK driven *Hprt* minigene, *pBT/PGK-Hprt.* Post transfection, cells were left overnight to settle in a non-selective medium. The growth media was changed 24 hours post electroporation to HAT medium to kill non-*pBT/PGK- Hprt* transfected cells and a subset of dishes were additionally exposed to a UV dose of $7.5Jm^{-2}$. Due to the ratio of DNA used in the transfection, cells that survived this selection were likely to also contain the *Ercc1* MG#13. Colonies were picked following this selection regime. Four of the colonies were expanded for use in UV survival studies. The UV survival studies were performed as with the fibroblasts but the plating density was lower in this experiment because of the rapid growth rate of the CHO cells; $2X10^4$ cells were plated/30mm petri dish and allowed to settle overnight to adhere to the dishes in a uniform layer. The following morning the cells were irradiated with various UV doses. The dishes were then cultured until the control dishes, which had not been UV irradiated, became confluent. The dishes were then fixed and stained for quantification of UV survival. The UV survival curves are shown in Figure 4.2.

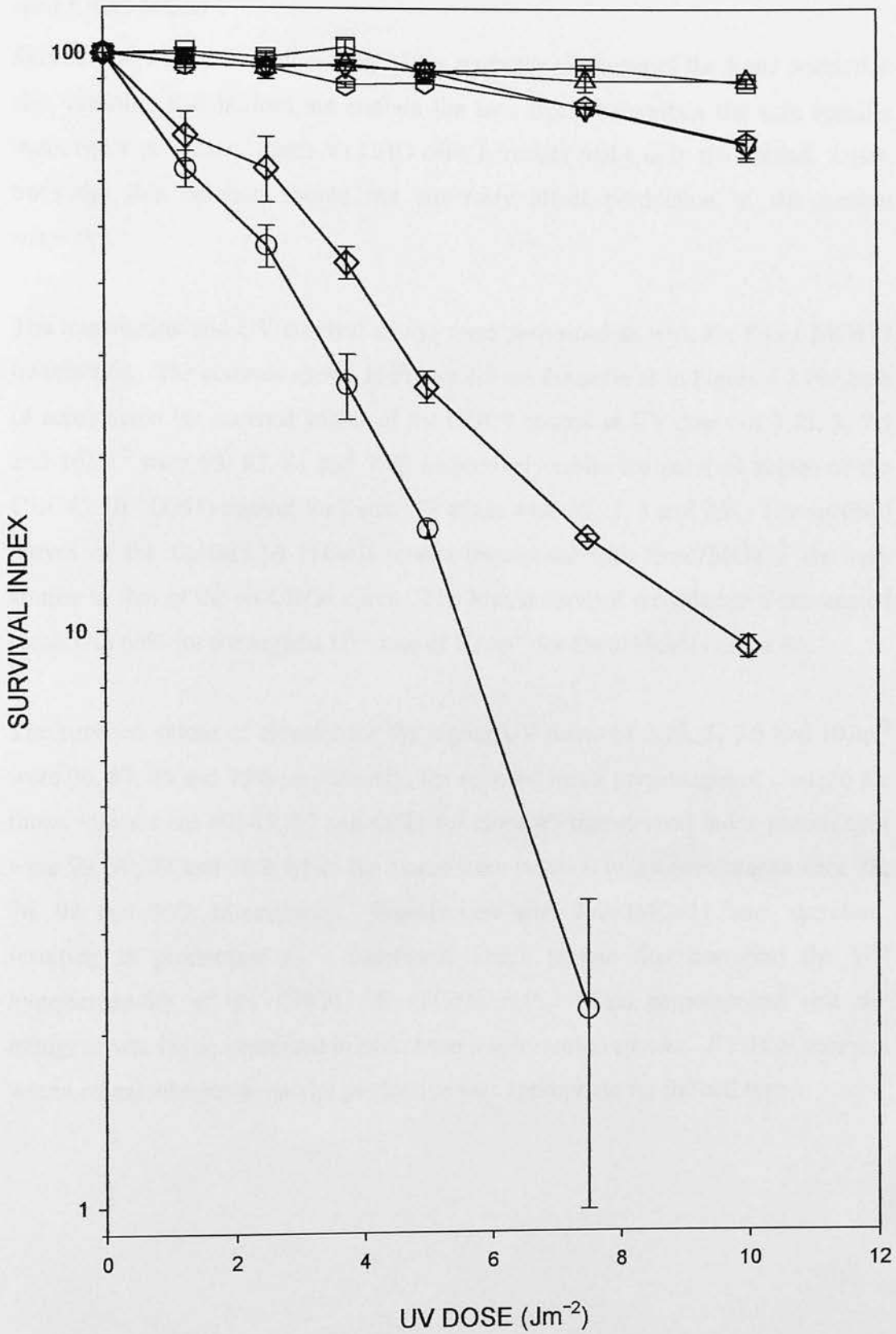
When the controls are examined the UV hypersensitivity of the non-transfected CHO43.3B (TG#1) control is apparent when the survival curve is compared to that of the wt CHO9. The percentage of cells surviving rapidly decreases with increasing UV dose in the non-transformed CHO43.3B (TG#1) sample, with survival index percentages of 63 and 47% for the low UV doses of 1.25 and $2.5Jm^{-2}$ dropping to 27, 15.2 and 2% for the higher doses of 3.75, 5, 7.5 and $10Jm^{-2}$. The D_{50} value for the

non-transformed CHO43.3B (TG#1), the dose at which only 50% of the cells survive, is merely 2.2Jm^{-2} while neither the wt CHO9 nor all bar one of the transfected clones drop as low as 50% survival for even the highest UV dose of 10J/m^{-2}. The survival curve of the wt CHO9 shows high resistance to UV with 95, 93, 81 and 70% survival for the highest UV doses of 3.75, 5, 7.5 and 10Jm^{-2}. Three of the *Ercc1* MG#13 transformed clones had survival curves which were very similar to that of the wt CHO9. For those respective UV doses, clone 10 had survival values of 88, 89, 81 and 70%; clone#4 had survival values of 103, 93, 95 and 89% and clone #9 had survival values of 96, 93, 90 and 90%. One clone, clone#3, had an unusual curve that lay between that of the wt and null controls. Survival values for the higher doses of 3.75, 5, 7.5 and 10Jm^{-2} were 44, 27, 15 and 9%, much lower than the other clones. There was partial correction of UV sensitivity by *Ercc1* MG#13 transfection with clone#3 but the other clones showed complete correction of UV hypersensitivity. However, transfection with *Ercc1* MG#13 generally rescued the UV hypersensitivity of the CHO43.3B (TG#1) cells to wild type levels and, therefore, resulted in production of functional protein in a cell line from a non-murine species. The RT-PCR results would reveal whether the unusual UV survival curve of clone #3 was due to unusual *Ercc1* expression.

In figure 4.2, the error bars of clone 3 do not overlap with the other clones. When tested using a two-tailed *t*-test, the difference between clone 3 and clone 4 at a UV dose of 10 Jm^{-2} is extremely significant (P=0.005). At the same UV dose, there is also an extremely significant difference between clone 3 and clone 9 (P=0.0001) and a very statistically significant difference between clone 3 and clone 10 (P= 0.0038). The error bars of clone 3 and the other clones do not overlap even at lower UV doses, indicating that clone 3 is atypical of the clones transfected with MG#13. When clone 3 is compared to untransfected cells at the UV dose of 7.5 Jm^{-2} there is a very statistically significant difference (P=0.0096), indicating that the clone has a greater resistance to UV than the untransfected cells at this dose, but the error bars almost overlap at the lower doses. Clone 3 was more like the untransfected cells that the transfected clones and this suggests that the minigene may not have integrated properly into this clone or was not being correctly expressed.

Figure 4.2 Correction of UV hypersensitivity of *Ercc1*-deficient CHO43.3B (TG#1) following transfection with *Ercc1* minigene#13

The UV survival curves for wild type (CHO9) and *Ercc1*-deficient CHO43.3B (TG#1) and four minigene-positive CHO43.3B (TG#1) clones are shown. The experiment was performed in duplicate and the mean value is shown for each point. Circle, non-transformed CHO43.3B (TG#1); inverted triangle, CHO9; diamond, clone#3; hexagon, clone#10; triangle, clone#9; square, clone#4.


100
10
1
SURVIVAL INDEX
0
2
4
6
8
10
12
UV DOSE (Jm-2)

4.3.2 Correction of UV hypersensitivity in CHO43.3B (TG#1) transformed with *Ercc1* MG#11

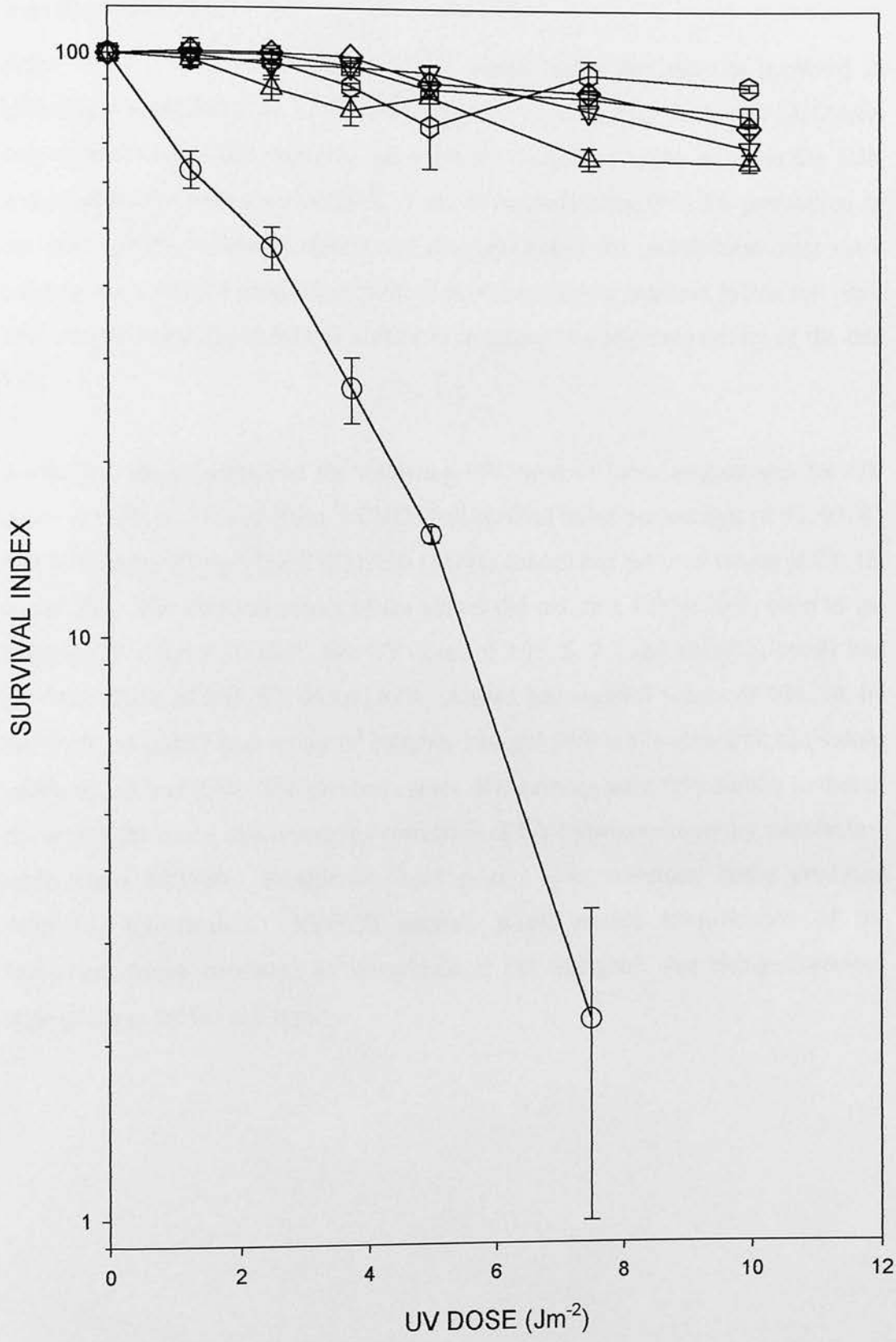
*Ercc1*MG#11 does not contain any of the sequence upstream of the *Kpn1* restriction site, meaning that is does not contain the area likely to contain the skin specific transcript's promoter. Since wt CHO cells normally make only the normal, 1.1kb, transcript this deletion should not adversely affect production of the normal transcript.

The transfection and UV survival assays were performed as with the *Ercc1* MG#13 transfection. The controls shown in Figure 4.3 are the same as in Figure 4.2 For ease of comparison the survival values of the CHO9 control at UV doses of 3.75, 5, 7.5 and 10Jm^{-2} were 95, 93, 81 and 70% respectively while the survival values of the CHO43.3B (TG#1) control for those UV doses were 27, 5, 2 and 2%. The survival curves of the CHO43.3B (TG#1) clones transfected with *Ercc1*MG#11 are very similar to that of the wt CHO9 curve. The lowest survival recorded in a transfected clone was 66% for the highest UV dose of 10 Jm^{-2} for *Ercc1*MG#11 clone #6.

The survival values of clone#2 for the higher UV doses of 3.75, 5, 7.5 and 10Jm^{-2} were 96, 87, 85 and 79% respectively; the survival index percentages of clone#6 for those values were 80, 85, 67 and 66%; for clone #5 the survival index percentages were 99, 90, 87 and 76% while for clone#9 the survival index percentages were 88, 76, 93 and 89% respectively. Transfection with *Ercc1*MG#11 was, therefore, resulting in production of a functional Ercc1 protein that corrected the UV hypersensitivity of the CHO43.3B (TG#1) cells. This demonstrated that the minigene was being expressed in cells from a non-murine species. RT-PCR analysis would reveal whether transcript production was appropriate for the cell type.

Figure 4.3 Correction of UV hypersensitivity of *Ercc1*-deficient CHO43.3B (TG#1) following transfection with *Ercc1* minigene#11

The UV survival curves for wild type (CHO9) and *Ercc1* deficient CHO43.3B (TG#1) and four minigene-positive CHO43.3B (TG#1) clones are shown. The experiment was performed in duplicate and the mean value is shown for each point. Circle, CHO43.3B (TG#1); inverted triangle, CHO9; diamond, clone#5; hexagon, clone#9; triangle, clone#6; square, clone#2.


100
10
1
SURVIVAL INDEX
0
2
4
6
8
10
12
UV DOSE (Jm-2)

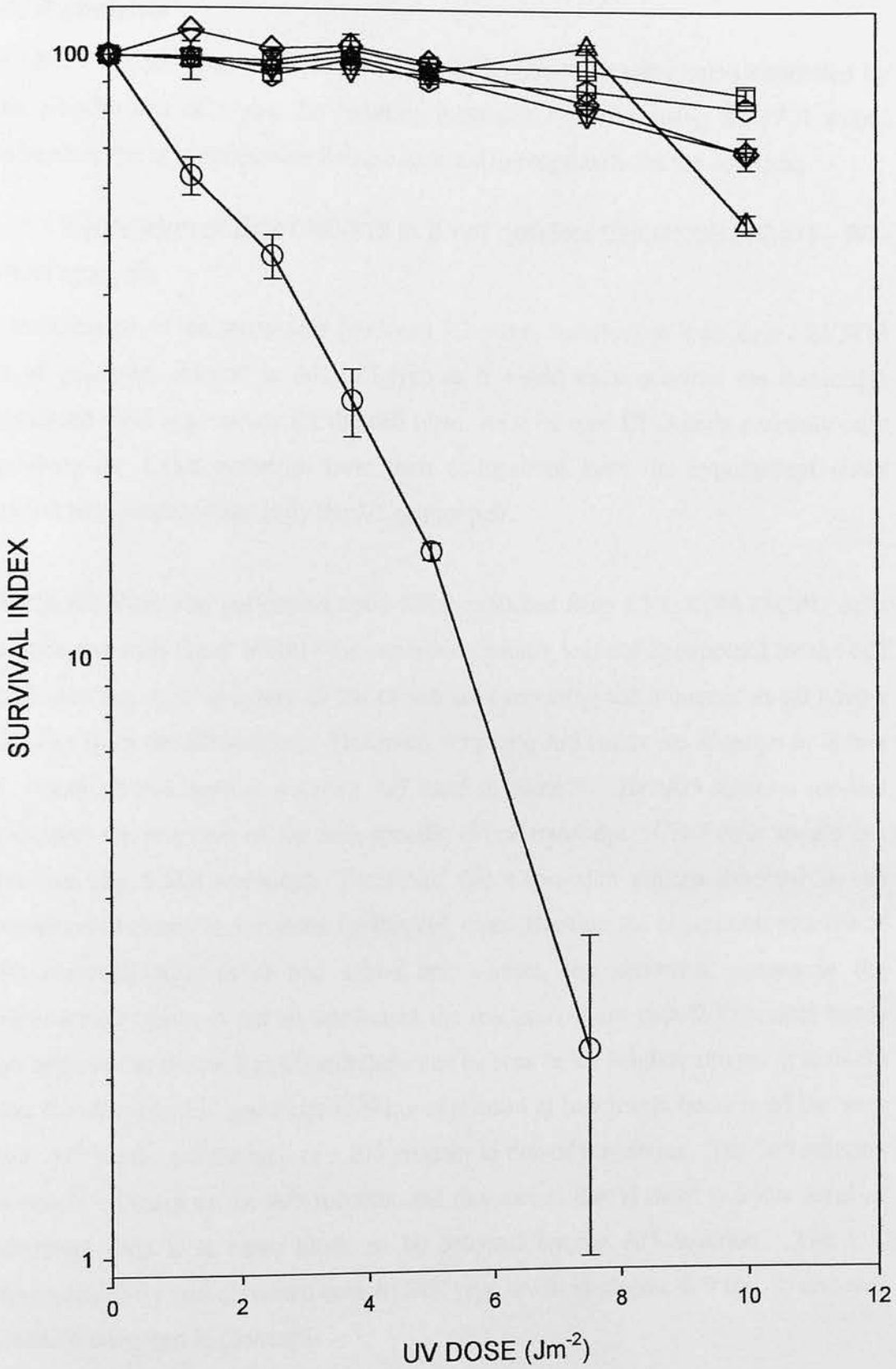
4.3.3 Correction of UV hypersensitivity in CHO43.3B (TG#1) transformed with *Ercc1* MG#20

Ercc1 MG#20 does not contain the (CT) repeat region that may be involved in initiating production of the 1.5kb skin specific *Ercc1* transcript. As the wt CHO cells only produce the 1.1kb transcript the same considerations apply as when the cells were transfected with *Ercc1*MG#11. Lack of sequences required for production of the skin specific transcript should not adversely effect the transformed cells since only the sequence for production of the normal transcript is required in this cell type. The normal transcript should be sufficient to correct the hypersensitivity of the cell line.

As before, the controls had the following UV survival index percentages for UV doses of 3.75, 5, 7.5 and 10Jm^{-2}. CHO9 had survival index percentages of 95, 93, 81 and 70% respectively. The CHO43.3B (TG#1) control had survival values of 27, 15, 2 and 2%. The survival values of the clones did not drop below 50%, even at the highest UV dose of 10 Jm^{-2}. For UV doses of 3.75, 5, 7.5 and 10Jm^{-2}, clone#1 had survival values of 103, 97, 96 and 87%, clone#1 had survival values of 104, 98, 84 and 70%, clone#17 had values of 100, 94, 104 and 53% while clone#18 had values of 99, 92, 88 and 83%. The survival curves of the clones were very similar to that of the wt CHO9 curve, demonstrating correction of UV hypersensitivity by transfection with *Ercc1* MG#20. Functional Ercc1 protein was, therefore, being produced following transfection. RT-PCR analysis would enable identification of the transcripts being produced to investigate if the minigene was being expressed appropriately for the cell type.

Figure 4.4 Correction of UV hypersensitivity of *Ercc1* deficient CHO43.3B (TG#1) following transfection with *Ercc1* minigene#20

The UV survival curves for wild type (CHO9) and *Ercc1*-deficient CHO43.3B (TG#1) and four minigene-positive CHO43.3B (TG#1) clones are shown. The experiment was performed in duplicate and the mean value is shown for each point. Circle, CHO43.3B (TG#1); inverted triangle, CHO9; diamond, clone#13; hexagon, clone#18; triangle, clone#14; square, clone#1.


100
10
1
SURVIVAL INDEX
0
2
4
6
8
10
12
UV DOSE (Jm-2)

4.4 Expression of minigenes in Ercc1 deficient CHO43.3B (TG#1) – RT-PCR analysis

RT-PCR analysis was required to show which transcripts were being expressed by the transformed cells and the resulting transcript detection using RT-PCR would indicate if the minigenes were being expressed appropriately for the cell type.

4.4.1 Expression of *Ercc1* MG#13 in *Ercc1* deficient CHO43.3B (TG#1) – RT-PCR analysis

Identification of the transcripts produced following transfection with *Ercc1* MG#13 is of particular interest in this cell type as it would show whether the transcripts produced were appropriate for the cell type. As wild type CHO cells normally only produce the 1.1kb transcript from their endogenous gene the hypothetical result would be a product from only the 2/5 primer pair.

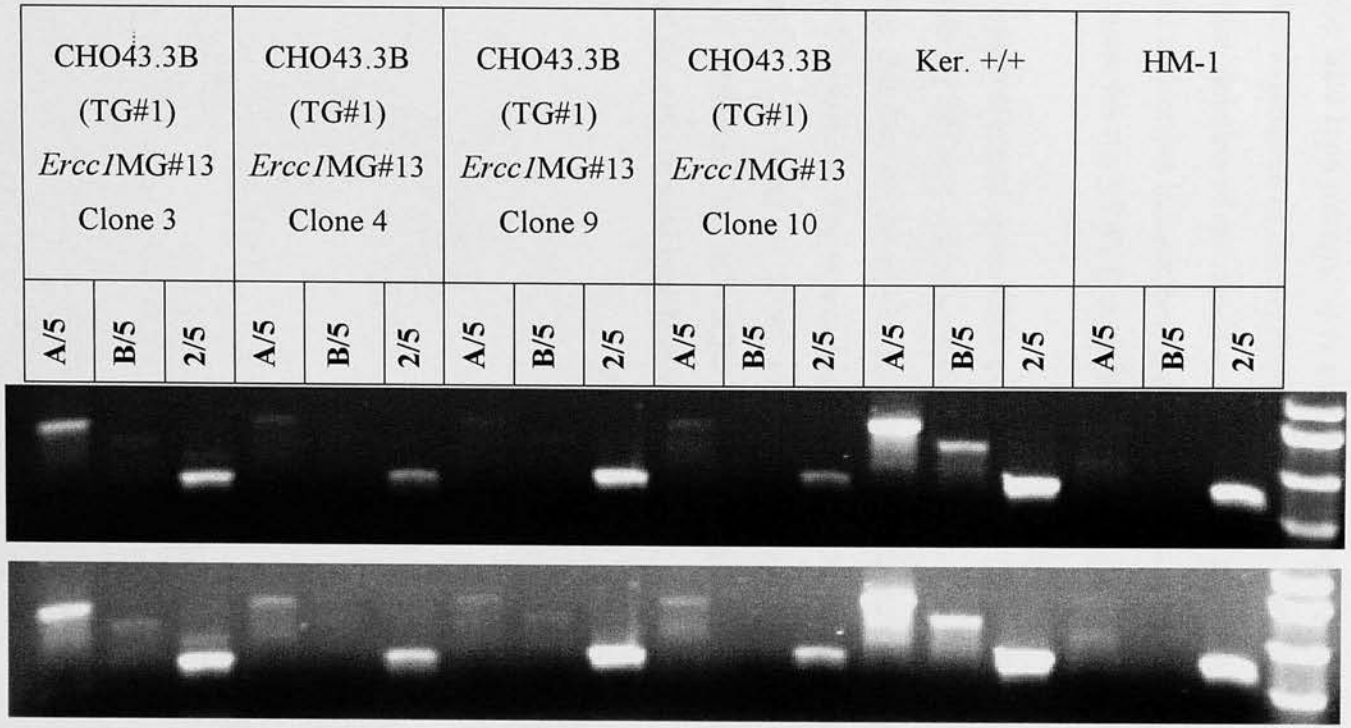
When RT-PCR was performed upon RNA extracted from CHO43.3B (TG#1) cells transfected with *Ercc1* MG#13 the expression pattern was not as expected for the cell type (see fig. 4.5). All four of the clones are expressing the minigene as all have a product from the 2/5 reaction. However, very pale A/5 bands are apparent in clones 4, 9 and 10 and there is a strong A/5 band in clone 3. The A/5 reaction product indicated the presence of the skin-specific *Ercc1* transcript. CHO cells should not produce the 1.5kb transcript. Therefore, this expression pattern detected in the transformed clones is abnormal for this cell type. Because the expression patterns of the controls, Ker. (+/+) and HM-1 are correct, the abnormal pattern in the transformed clones is not an artefact of the reaction. Very pale B/5 product bands are apparent in clones 3 and 9 and these can be seen in the brighter image. It is likely that the skin-specific transcript is being expressed at low levels because of the very pale A/5 bands and the lack of a B/5 product in two of the clones. The B/5 reaction is not as efficient as the A/5 reaction and this means that if there is a low level of transcript then it is more likely to be detected by the A/5 reaction. The UV hypersensitivity was corrected back to wild type levels in clones, 4, 9 and 10 and was partially corrected in clones#3.

Figure 4.5 Expression of *Ercc1* MG#13 in CHO43.3B (TG#1) - RT-PCR

RT-PCR analysis of RNA extracted from *Ercc1*-deficient CHO43.3B (TG#1) clones transformed with *Ercc1*MG#13. Total RNA was prepared and reverse transcribed using oligo-dT primers. The *Ercc1* cDNA was then PCR amplified from the cDNA pool of each clone. Three PCR reactions were performed on cDNA from each clone using the following forward primers respectively; primer 'A' is specific a region upstream of the normal transcriptional start site, primer 'B' is specific to a region upstream of the *Kpn1* restriction site and primer '2' is specific to exon 2. Primer '5' is specific to a region in exon 5 and is the reverse primer for each of the three reactions.

20μg of cDNA was used in each PCR reaction. The PCR was set to 30 cycles for all three primer pairs. The PCR products were electrophoresed at 100V on an ethidium bromide/1% agarose gel for 1.5 hours.

The lower panel is a higher exposure of the gel shown in the upper panel and more clearly shows the presence of a PCR product in the A/5 reactions of clones 4, 9 and 10. The PCR product of the B/5 reaction is seen in clones 3 and 9.


CHO43.3B (TG#1) *Ercc1*MG#13 Clone 3
CHO43.3B (TG#1) *Ercc1*MG#13 Clone 4
CHO43.3B (TG#1) *Ercc1*MG#13 Clone 9
CHO43.3B (TG#1) *Ercc1*MG#13 Clone 10
Ker. +/+
HM-1
A/5
B/5
2/5
A/5
B/5
2/5
A/5
B/5
2/5
A/5
B/5
2/5
A/5
B/5
2/5
A/5
B/5
2/5

4.4.2 Expression of *Ercc1* MG#11 in *Ercc1*-deficient CHO43.3B (TG#1) – RT-PCR analysis

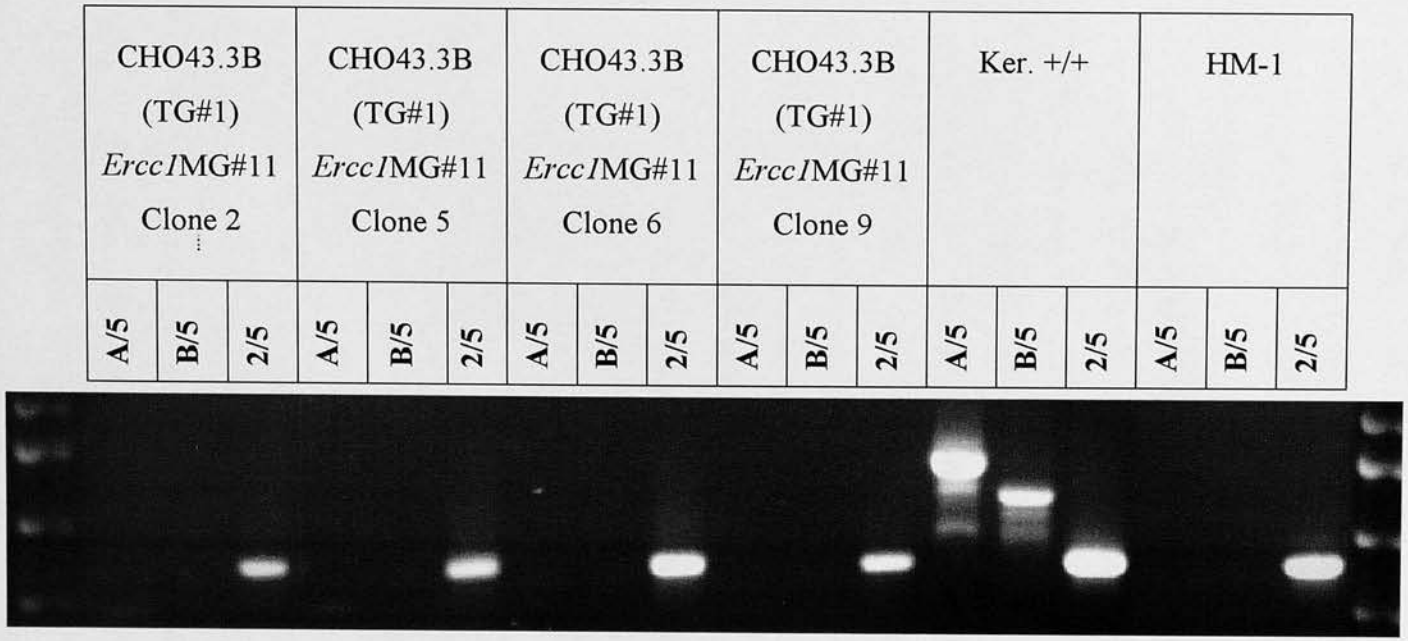
Because of the deletion in *Ercc1*MG#11 it is not possible for transfection to result in production of the 1.5kb transcript. As CHO cells normally only produce the 1.1kb transcript, expression of the minigene in this cell type should result in a situation which matches the wild type pattern. It is possible that the region deleted may be required for correct expression of the normal transcript as it could have a controlling function. Therefore, deletion of the region upstream of the *Kpn1* site could also lead to abnormal *Ercc1* expression in transformed cells. If the minigene is functional then the hypothetical result for the RT-PCR is product from only the 2/5 reaction.

When RT-PCR was performed upon RNA extracted from CHO43.3B (TG#1) cells transfected with *Ercc1* MG#11 the expression pattern was as expected for the cell type (see fig. 4.6). All four of the clones are expressing the minigene as all have a product from the 2/5 reaction. The expression patterns of the controls, Ker. +/+ and HM-1, are correct. The level of *Ercc1* expression following transfection with *Ercc1*MG#11 and detected by RT-PCR was sufficient to correct the UV hypersensitivity of the cells to wild type levels (see fig. 4.3).

Figure 4.6 Expression of *Ercc1* MG#11 in CHO43.3B (TG#1) - RT-PCR

RT-PCR analysis of RNA extracted from *Ercc1*-deficient CHO43.3B (TG#1) clones transformed with *Ercc1*MG#11. Total RNA was prepared and reverse transcribed using oligo-dT primers. The *Ercc1* cDNA was then PCR amplified from the cDNA pool of each clone. Three PCR reactions were performed on cDNA from each clone using the following forward primers respectively; primer 'A' is specific a region upstream of the normal transcriptional start site, primer 'B' is specific to a region upstream of the *Kpn1* restriction site and primer '2' is specific to exon 2. Primer '5' is specific to a region in exon 5 and is the reverse primer for each of the three reactions.

20μg of cDNA was used in each PCR reaction. The PCR was set to 30 cycles for all three primer pairs. The PCR products were electrophoresed at 100V on an ethidium bromide/1% agarose gel for 1.5 hours.


CHO43.3B (TG#1) *Ercc1*MG#11 Clone 2
CHO43.3B (TG#1) *Ercc1*MG#11 Clone 5
CHO43.3B (TG#1) *Ercc1*MG#11 Clone 6
CHO43.3B (TG#1) *Ercc1*MG#11 Clone 9
Ker. +/+
HM-1
A/5
B/5
2/5
A/5
B/5
2/5
A/5
B/5
2/5
A/5
B/5
2/5
A/5
B/5
2/5
A/5
B/5
2/5

4.4.3 Expression of *Ercc1* MG#20 in *Ercc1*-deficient CHO43.3B (TG#1) – RT-PCR analysis

Identification of the transcripts produced following transfection with *Ercc1* MG#20 is interesting because the deleted CT repeat region may be important for correct expression of *Ercc1*. As CHO cells normally only produce the 1.1kb transcript the hypothetical result would be a product from only the 2/5 primer pair. However, if the (CT) repeat region is required for control of 1.5kb transcript production then it is possible that abnormal production of transcripts may be detected following transfection with this minigene.

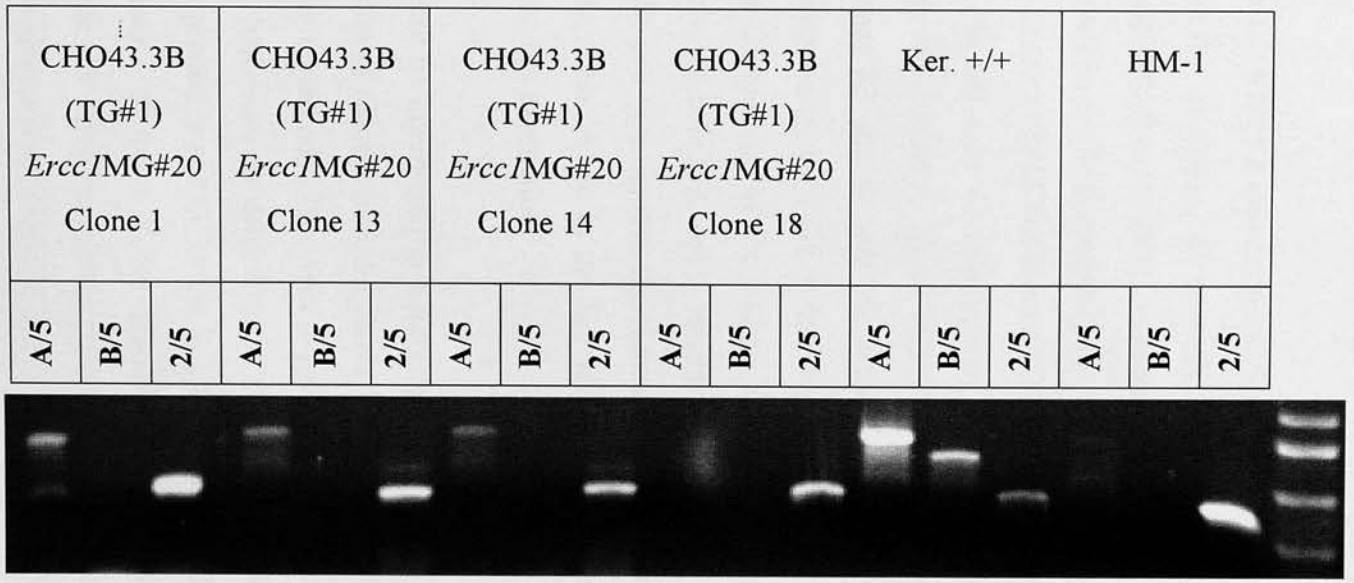
When RT-PCR was performed upon RNA extracted from CHO43.3B (TG#1) cells transfected with *Ercc1* MG#20 the expression pattern was not as expected for the cell type in 3 of the 4 clones (see fig. 4.7). All four of the clones are expressing the minigene as all have a product from the 2/5 reaction. However, pale A/5 bands are apparent in clones #1, #13 and #14. The A/5 reaction product indicated the presence of the skin-specific *Ercc1* transcript. Therefore, this expression pattern detected in clones #1, #13 and #14 of the cells is abnormal for this cell type. This aberrant expression pattern was also observed in the clones transfected with *Ercc1* MG#13. CHO cells should not produce the 1.5kb transcript. Because the expression patterns of the controls, Ker. +/+ and HM-1, are correct the abnormal pattern in the transformed clones is not an artefact of the reaction. There are no B/5 product bands. It is likely that the skin-specific transcript is being expressed at low levels because of the pale A/5 bands and the lack of a B/5 product for those clones which have A/5 bands. The B/5 reaction is not as efficient as the A/5 reaction and this means that if there is a low level of transcript then it is more likely to be detected by the A/5 reaction.

The clones with A/5 bands, clone#1, #13 and #14 had UV survival curves which clustered around the wild type pattern and the hypersensitivity of clone #18, which had the correct pattern of expression for the cell type, was as effectively corrected by the MG expression as the three clones which had abnormal 1.5kb transcript production. This suggests that the expression of the 1.5k transcript did not confer a survival advantage upon the clones in comparison to the clone with only the 1.1kb transcript.

Figure 4.7 Expression of *Ercc1* MG#20 in CHO43.3B (TG#1) - RT-PCR

RT-PCR analysis of RNA extracted from *Ercc1*-deficient CHO43.3B (TG#1) clones transformed with *Ercc1*MG#20. Total RNA was prepared and reverse transcribed using oligo-dT primers. The *Ercc1* cDNA was then PCR amplified from the cDNA pool of each clone. Three PCR reactions were performed on cDNA from each clone using the following forward primers respectively; primer 'A' is specific a region upstream of the normal transcriptional start site, primer 'B' is specific to a region upstream of the *Kpn1* restriction site and primer '2' is specific to exon 2. Primer '5' is specific to a region in exon 5 and is the reverse primer for each of the three reactions.

20μg of cDNA was used in each PCR reaction. The PCR was set to 30 cycles for all three primer pairs. The PCR products were electrophoresed at 100V on an ethidium bromide/1% agarose gel for 1.5 hours.


CHO43.3B (TG#1) *Ercc1*MG#20 Clone 1
CHO43.3B (TG#1) *Ercc1*MG#20 Clone 13
CHO43.3B (TG#1) *Ercc1*MG#20 Clone 14
CHO43.3B (TG#1) *Ercc1*MG#20 Clone 18
Ker. +/+
HM-1
A/5
B/5
2/5
A/5
B/5
2/5
A/5
B/5
2/5
A/5
B/5
2/5
A/5
B/5
2/5
A/5
B/5
2/5

4.5 Conclusions

Transfection with *Ercc1 minigenes* #13, #11 and #20 resulted in correction of the UV hypersensitivity of the CHO43.3B (TG#1) cell line and, by definition, in production of functional Ercc1 protein in cells from a non-murine species. However, RT-PCR analysis revealed that the expression patterns of the minigenes were variable.

In the case of *Ercc1*MG#11 it was not possible to produce skin transcripts as the skin TSS is missing in this minigene. As expected, only the 1.1kb transcript was present in cells transfected with *Ercc1*MG#11. Correct expression of *Ercc1* MG#13 and #20 would have resulted in production of only the 1.1kb transcript. However, the clones transfected with *Ercc1* MG#13 had abnormal *Ercc1* transcript production for the cell type (fig. 4.5). In the case of transfection with MG#13 expression of the skin specific transcript was variable. *Ercc1* MG#13 was correctly expressed in clone #9 However, in clones #3 and #4 and #9, cells that should only express the 1.1kb transcript were expressing both the normal and skin transcripts, the band intensity was very low in comparison to the control Ker. (+/+). There is, therefore, some semblance of the transfected cells retaining specificity of *Ercc1* expression appropriate for the cell type since in the clones expressing skin=specific transcript in most cases product was absent in the B/5 reaction, indicating likely low levels of skin-specific transcript expression.

These results could be interpreted in a number of ways. Firstly, the minigenes did result in production of functional Ercc1 as the UV hypersensitivity of the CHO43.3B (TG#1) cells was corrected. This was a promising result as it indicated that the minigenes were functional in cells from a non-murine species. However, 1.5kb transcripts were produced at low levels in all four of the clones transformed with *Ercc1* MG#13 and 3 of the 4 clones transformed with *Ercc1* MG#20. It was a possibility that the UV selection regime following transformation could be resulting in a bias towards clones with abnormal *Ercc1* expression. Perhaps abnormal *Ercc1* expression, the presence of 1.5kb transcripts, was required to survive the UV selection regime. It was a possibility that had to be explored so alterations were made to the experimental technique in the following keratinocyte transformations.

Clones were selected having been co-selected with HAT and UV as in previous experiments but in these experiments only HAT selection was applied to some dishes following electroporation to discover if there was any difference between those selected by either method at the UV survival and RT-PCR analysis.

In addition, it was possible that the clones being selected were not representative of the entire population of transformed cells. To address this possibility, cells were plated following electroporation and the transformed cells were treated as pools as well as selecting individual colonies for study.

Chapter 5

Use of *Ercc1* minigenes to correct UV hypersensitivity of *Ercc1*-deficient keratinocytes

5.1 Expression of minigenes in cells which normally have a high skin-specific to normal transcript ratio

Transformation of the *Ercc1*-deficient fibroblasts demonstrated that the minigenes were functional in a murine cell line that normally expressed both transcripts. However, the fibroblast pattern of *Ercc1* expression, although similar to that of skin, is not identical. The murine skin *Ercc1* expression pattern has a high skin specific (1.5kb) to normal (1.1kb) *Ercc1* transcript ratio. Fortunately a suitable cell line was available for further investigation of the importance of various sequences to the control of the skin specific expression pattern. Dr. Selfridge (Melton Lab.) had demonstrated by northern blot that cultured wild type murine keratinocytes have an identical *Ercc1* expression pattern to that of skin (fig. 4.1). *Ercc1*-deficient murine keratinocytes had previously been isolated and were also *Hprt*-deficient. This made them ideal candidates for use in the minigene transfection experiments as the *Hprt* status of the cells meant that the same co-transfection method could be performed as in the fibroblast and CHO experiments.

Keratinocytes are technically more demanding to culture than the other cell types studied. This meant that some changes were required at the stage where transformed clones were selected because these cells do not grow well at low densities. The cells also require specially treated growth media and trypsinisation conditions. The slow growth rate of the cells increases the time spent expanding cell numbers and this increase the risk of fungal contamination. Keratinocytes originate from the epidermal layer and are more resistant to UV than fibroblasts and CHO cells. This meant that higher UV doses were required in the UV survival experiments before any lethal effects were apparent.

To further investigate the possibility that the UV selection regime following transfection could be favouring clones with abnormal minigene expression a second component was added to the UV survival experiments. As well as subjecting transformed clones to selection by HAT and UV prior to the UV survival experiments, a subset of clones had HAT selection only. In tandem with this, for each of the two selection conditions mentioned, pools of cells were subjected to the

same treatment. These pools were to be representative of entire transformed cell populations. Differences between the results for the UV survival and RT-PCR with HAT versus HAT + UV selection and for pools versus clones would assist in determining if the UV selection following transformation was causing clones expressing the minigenes abnormally to be favoured.

Following the UV survival experiments with the clones and pools, pools which had undergone HAT only selection following transfection were compared to pools which had undergone HAT + UV selection by means of RT-PCR. The *Ercc1* MG#20 pool which had been subjected to HAT only selection was then repeatedly irradiated with UV to find if this would induce a change towards the *Ercc1* expression pattern of the pools which had undergone UV and HAT selection following transfection.

5.2 Number of colonies comprising keratinocyte pools

The number of colonies in each pool of transformed cells in shown in table 5.1. For each minigene there were two pools that had either undergone selection by HAT only following electroporation, or HAT and UV selection. The number of colonies was counted before the cells were trypsinised from the original plating dishes to be expanded in flasks for UV survival assays and RT-PCR. For each minigene the co-selection results in fewer colonies surviving than the HAT selection alone. There was a large difference between the ratio of HAT+UV selected colonies to HAT selected colonies between pools transformed with *Ercc1* MG#13 and #20 when compared to the ratio seen in cells transformed with *Ercc1*MG#11. The ratio of HAT + UV selected colonies to HAT selected colonies in the MG#11 pool was only 0.483, demonstrating that transfection with *Ercc1*MG#11 did not correct the UV hypersensitivity of the cells as well as transfection with *Ercc1* MG#13 and #20.

For minigene 13 there were 163 fewer colonies in the HAT+UV pool (total of 501 colonies) than the HAT only pool (664 colonies). In the cells transformed with *Ercc1*MG#11 there were 513 fewer colonies in the HAT+UV pool (554 colonies) than the HAT only pool (1057 colonies). In the cells transformed with *Ercc1*MG#20 there were 222 fewer colonies in the HAT+UV pool (976 colonies) than the HAT

only pool (1198 colonies). The difference between the two selection regimes is an indication of how often the *pBT/PGK-Hprt* minigene alone transforms a cell during electroporation in comparison to how often both the *pBT/PGK-Hprt* minigene and an *Ercc1* minigene are actually incorporated. This is because selection with HAT removes cells that do not contain the *pBT/PGK-Hprt* minigene while co-selection with HAT and UV removes cells which do not contain both *pBT/PGK-Hprt* and an *Ercc1* minigene.

Table 5.1 Colony counts following *Ercc1*-minigene-transfection and HAT/UV or HAT only selection in *Ercc1*-deficient keratinocyte pools.

For each minigene there are two pools – one that underwent HAT selection and one that was HAT+UV. The plating density was 3.75×10^5 cells per dish. In each row of the table, details of the *Ercc1* MG, selection regime and the total number of colonies comprising the pool is shown.

Electroporation DNA	**No. colonies in HAT selected pool (n=1)**	**No. colonies in HAT +UV selected pool (n=1)**	**Ratio of HAT + UV selected colonies to HAT selected colonies**
Ercc1 MG#13 + pBT/PGK-Hprt	664	501	0.755
Ercc1 MG#11 + pBT/PGK-Hprt	1057	511	0.483
Ercc1 MG#20 + pBT/PGK-Hprt	1198	976	0.815

5.3 Correction of UV hypersensitivity of *Ercc1*-deficient murine keratinocytes with *Ercc1* minigenes

As mentioned earlier, the pattern of *Ercc1* expression that is observed in wild type murine keratinocytes is identical to that of skin, as keratinocytes have a high skin specific to normal transcript ratio. Hypothetically, transfection of *Ercc1*-deficient murine keratinocytes with *Ercc1* MG#13 should result in the wild type pattern of expression if the minigene is fully functional in the cultured keratinocytes. As in the previous experiments, the *Ercc1*-deficient keratinocytes were co-transfected with *Ercc1* MG#13 and *pBT/PGK-Hprt* at a 3:1 DNA ratio. The vector DNA had been linearised with *NotI* prior to transfection. Following transfection the cells were left to settle and adhere to the surface of the petri dishes overnight. 24 hours post electroporation the medium was replaced with selective HAT medium to prevent growth of cells that had not been transfected with the *pBT/PGK-Hprt* minigene. Due to the 3:1 ratio of *Ercc1* to *pBT/PGK-Hprt* DNA the cells that survived this selection were likely to also contain the *Ercc1* MG#13. At this stage half of the dishes also underwent UV selection. Clones were picked from dishes that had had HAT only or HAT+UV selection and entire pools of transformed cells were also cultured from dishes that had had either treatment. These pools and clones (following expansion) were then used in UV survival assays and the minigene expression patterns were investigated using RT-PCR.

5.3.1 Correction of UV hypersensitivity of Ercc1-deficient murine keratinocyte clones following transfection with Ercc1 MG#13

Assuming that *Ercc1* MG#13 has all the sequence elements for correct expression of*Ercc1,* transformation of *Ercc1* null keratinocytes should lead to correction of UV hypersensitivity and an expression pattern that matches that of wild type keratinocytes.

Following electroporation, five colonies from two selection regimes were expanded for use in UV survival studies. $4X10^4$ cells were plated/30mm petri dish and allowed to settle overnight, to adhere to the dishes in a uniform layer. The following morning the cells were irradiated with various UV doses. The dishes were then cultured until

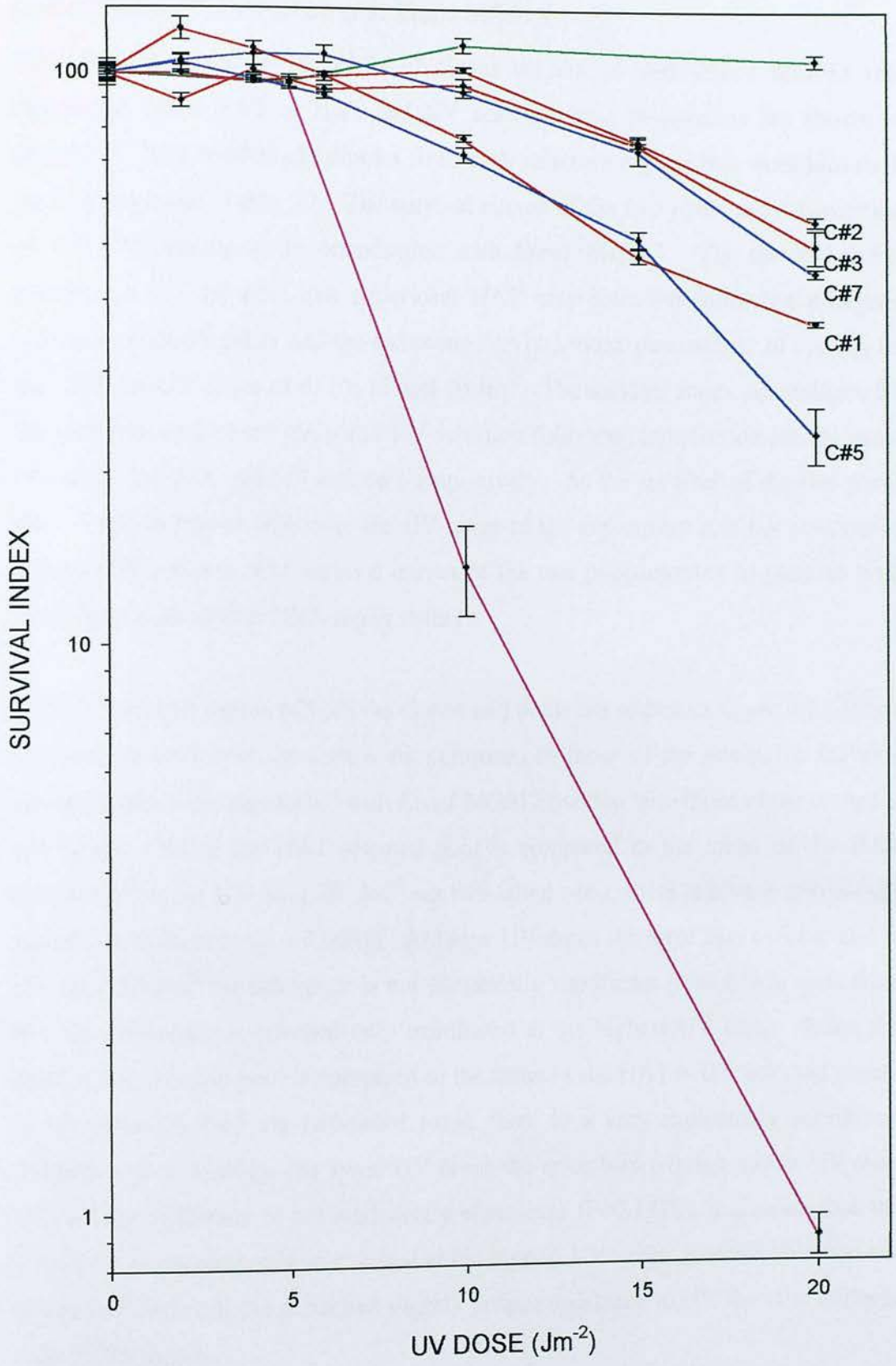
the control dishes, which had not been irradiated, became confluent. The dishes were then fixed and stained for quantification of survival. The UV survival curves for wild type (Ker. +/+), *Ercc1*-deficient (Ker. -/-) murine keratinocytes, three clones which had HAT selection only following transfection (clones 1, 2, 3) and two clones which had HAT and UV selection (clones 5 and 7) are shown in figure 5.1.

When the controls are examined the UV hypersensitivity of the *Ercc1*-deficient keratinocyte (Ker.(-/-)) control is obvious when compared to that of the wild type keratinocytes (Ker.(+/+)) . The percentage of cells surviving rapidly decreases in the Ker.(-/-) sample, with survival index percentages of 96 and 14% for the low UV doses of 5 and 10Jm^{-2} dropping to 1% for the higher dose of 20Jm^{-2} . The D_{50} value for the Ker.(-/-), the dose at which only 50% of the cells survive, is 6.5 Jm^{-2}. In contrast, the wild type keratinocytes show the expected resistance to UV irradiation, with 97 and 113% survival at the lower UV doses of 5 and 10Jm^{-2} and 106% survival at 20 Jm^{-2}. The survival index score does not drop as low as 50% over the UV dose range of the experiment. All the clones transfected with *Ercc1* MG#13 show corrected UV sensitivity with the survival curves shifted towards the wild type curve. The clones that underwent HAT selection only following transfection had the following survival index percentages for the UV doses of 6, 10, 15 and 20 Jm^{-2}; *Ercc1* MG#13 clone 1 had 100, 77, 48 and 37% respectively (D_{50} value13 Jm^{-2}), *Ercc1* MG#13 clone 2 survival index percentages for those doses were 99, 102, 76 and 55% (D_{50} value 19 Jm^{-2}) and *Ercc1* MG#13 clone 3 had survival index percentages of 94, 96, 75 and 50% (D_{50} value18 Jm^{-2}). The mean survival of the set of clones selected with HAT is similar to that of the set of clones selected with HAT+UV, indicating little difference in the UV resistance of the clones resulting from the two selection regimes. The survival index percentages of the HAT + UV selected *Ercc1* MG#13 clone 7 for UV doses of 6, 10, 15 and 20 Jm^{-2} are 108, 93, 74 and 45% respectively (D_{50} value 17.5 Jm^{-2}), while those of *Ercc1* MG#13 clone 5 are 92, 72, 51 and 23% respectively (D_{50} value13 Jm^{-2}). The average D_{50} value for the clones which were HAT only selected following transfection is 16.7 Jm^{-2} while the average D_{50} value for the clones that underwent HAT and UV selection post transfection is 15.3 Jm^{-2}.

These survival curves show that transfection with *Ercc1* MG#13 corrected the UV hypersensitivity of the keratinocytes and that selection with HAT + UV following transfection did not lead to selection of clones with a greater ability to repair UV damage than clones which had not undergone UV selection post transfection. RT-PCR analysis was required to investigate which transcripts the transfected clones were producing.

Figure 5.1 Correction of UV sensitivity of clones of *Ercc1*-null keratinocytes (Ker.(-/-) following transfection with *Ercc1* MG#13

The UV survival curves for wild type (Ker. +/+), *Ercc1*-deficient (Ker.-/-) murine keratinocytes, three clones which had HAT selection only following transfection (clones 1, 2, 3 – in red) and two clones which had HAT and UV selection (clones 5 and 7 – in blue) are shown. Pink, Ker.(-/-); green Ker.(+/+); red, clones from HAT only selection regime; blue, clones from HAT+UV selection regime. Each UV point represents the mean of two separate experiments.


100
10
1
SURVIVAL INDEX
0
5
10
15
20
UV DOSE (Jm-2)
C#2
C#3
C#7
C#1
C#5

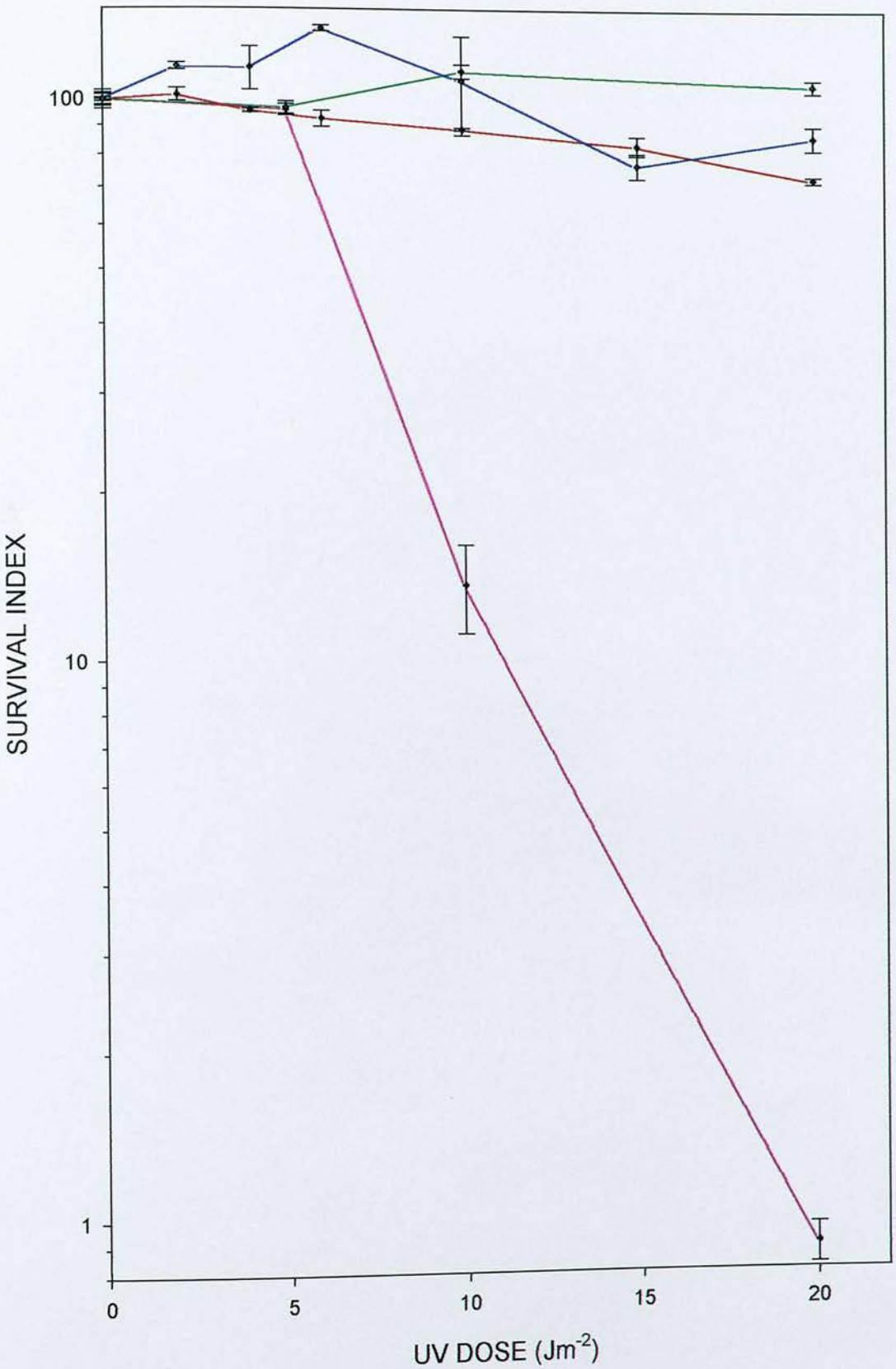
5.3.2 Correction of UV hypersensitivity of *Ercc1*-deficient murine keratinocyte pools following transfection with *Ercc1* MG#13

The survival curves of the pools of clones transfected with *Ercc1* MG#13 that underwent either HAT or HAT and UV selection post transfection are shown in figure 5.2. The number of colonies from each selection regime that went into each pool are shown in Table 5.1. The survival curves of the two pools show correction of UV hypersensitivity by transfection with *Ercc1* MG#13. The survival index percentages for the pool that underwent HAT only selection following minigene transfection (HAT pool) had the following survival index percentages of 93, 89, 83 and 72% for UV doses of 6, 10, 15 and 20 Jm^{-2}. The survival index percentages for the pool that underwent HAT and UV selection following transfection for the same UV doses are 134, 108, 77 and 88% respectively. As the survival of the two pools did not fall as low as 50% over the UV range of the experiment it is not possible to calculate D_{50} values. The survival curves of the two pools overlap in parts so both pools have very similar DNA repair ability.

The UV survival curves of both the clones and pools are shown in figure 5.3. When the survival curves of the clones are compared to those of the pools, the survival curves of the pools transfected with *Ercc1* MG#13 overlap with those of the clones at low doses. When the HAT selected pool is compared to the mean of the HAT selected clones at UV dose 20 Jm^{-2} via two-tailed *t*-test, there is a very statistically significant difference ($P = 0.0085$). At lower UV doses the error bars overlap and at UV dose 10 Jm^{-2} the difference is not statistically significant ($P=0.8090$), indicating that the difference in survival only manifested at the highest UV dose. When the HAT + UV selected pool is compared to the mean of the HAT + UV selected clones at UV dose 20 Jm^{-2} via two-tailed *t*-test, there is a very statistically significant difference ($P = 0.0042$). At lower UV doses the error bars overlap and at UV dose 10 Jm^{-2} the difference is not statistically significant ($P=0.1778$), indicating that the difference in survival only manifested at the highest UV dose. It is possible that the mixed population of the pools had slightly greater resistance to UV than the uniform cells of the clones.

Figure 5.2 Correction of UV sensitivity of pools of *Ercc1*-null keratinocytes Ker.(-/-) following transfection with *Ercc1* MG#13

The UV survival curves for wild type (Ker. +/+), Ercc1-deficient (Ker.-/-) murine keratinocytes, one pool which had HAT selection only following transfection with MG#13 (pool HAT) and one pool which had HAT and UV selection (pool HAT + UV) are shown. Each UV point represents the mean of two separate experiments. Pink, Ker.(-/-); green, Ker.(+/+); blue, HAT + UV selected pool; red, HAT selected pool.


100
10
1
SURVIVAL INDEX
0
5
10
15
20
UV DOSE (Jm-2)

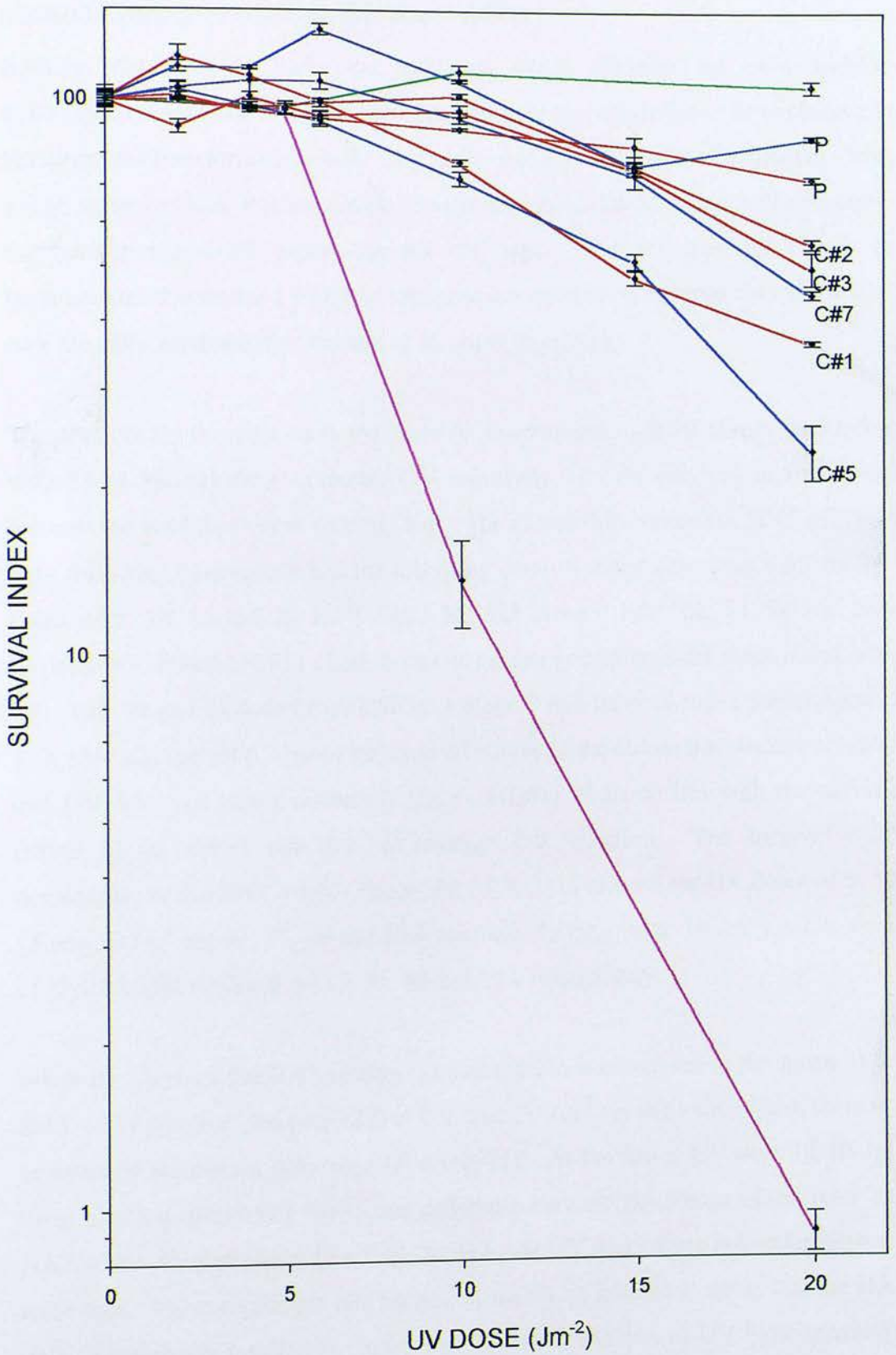

100
10
1
SURVIVAL INDEX
0
5
10
15
20
UV DOSE (Jm-2)
P
P
C#2
C#3
C#7
C#1
C#5

Figure 5.3 Correction of UV sensitivity of clones and pools of *Ercc1*-null keratinocytes Ker.(-/-) following transfection with *Ercc1* MG#13

The UV survival curves for wild type (Ker. +/+), *Ercc1*-deficient (Ker.-/-) murine keratinocytes, one pool plus three clones (clones, 1, 2 and 3) which had HAT selection only following transfection with MG#13 are shown in red. One pool plus 2 clones (clones 5 and 7) which had HAT and UV selection are shown in blue. Each UV point represents the mean of two separate experiments. Pink, Ker.(-/-); green, Ker.(+/+); blue, HAT + UV selected pool and clones; red, HAT selected pool and clones.

5.3.3 Correction of UV hypersensitivity of *Ercc1*-deficient murine keratinocyte clones following transfection with *Ercc1* MG#11

Because *Ercc1*MG#11 lacks the sequence which contains the skin specific transcriptional start site it should not be possible for this transcript to be expressed in keratinocytes transformed with the minigene. Because wild type keratinocytes have a high upper to lower transcript ratio, the expression of this minigene will not mimic the normal expression pattern for the cell type. The UV survival curves of keratinocytes transformed with this minigene are interesting because they show how well the cells survive in the absence of an upper transcript.

The controls are the same as in the previous experiments. All the clones transfected with *Ercc1* MG#11 show corrected UV sensitivity with the survival curves shifted towards the wild type curve (see fig. 5.4). The clones that underwent HAT selection only following transfection had the following survival index percentages for the UV doses of 6, 10, 15 and 20 Jm^{-2}; *Ercc1* MG#11 clone 2 had 106, 95, 84 and 56% respectively, *Ercc1* MG#11 clone 3 survival index percentages for those doses were 114, 103, 94 and 81% and *Ercc1* MG#11 clone 7 had survival index percentages of 133, 136, 104 and 98%. One of the survival curves of the clones that underwent HAT and UV selection post transfection (*Ercc1* MG#11 clone 5) lies with the survival curves of the clones that did not undergo UV selection. The survival index percentages of the HAT + UV selected *Ercc1* MG#11 clone 4 for UV doses of 6, 10, 15 and 20 Jm^{-2} are 97, 77, 60 and 29% respectively (D_{50} value 16 Jm^{-2}), while those of *Ercc1* MG#11 clone 5 are 99, 94, 95 and 79% respectively.

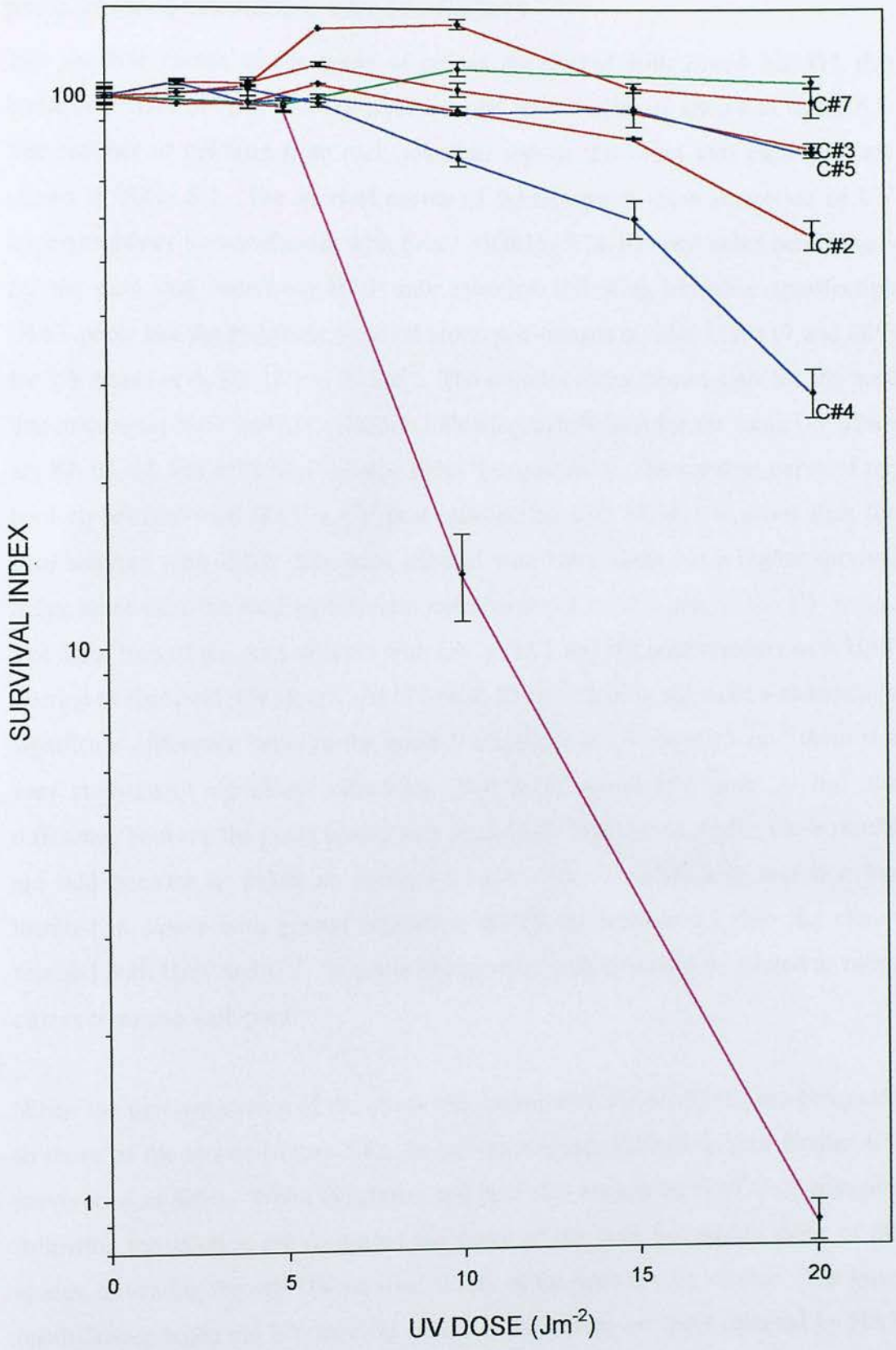
When the mean of the HAT selected clones (78.25) is compared to the mean of the HAT + UV selected clones (53.83) at UV dose 20 Jm^{-2} via two-tailed *t*-test, there is a statistically significant difference ($P = 0.0371$). At the lower UV dose of 10 Jm^{-2} there is still a statistically significant difference between the means of the HAT and HAT + UV selected clones ($P = 0.312$). At lower UV doses there is overlap between error bars. These results are odd because it makes no biological sense that the HAT only selection has resulted in clones with greater correction of UV hypersensitivity than the clones selected with HAT and UV. Perhaps there was variability in the

correction of UV hypersensitivity between clones in the two selection regime and these differences represent a sampling effect. Without testing a larger number of clones it is not possible to know how representative the clones investigated here were.

These survival curves show that transfection with *Ercc1* MG#11 corrected the UV hypersensitivity of the *Ercc1/Hprt* null keratinocytes and that selection with HAT + UV following transfection did not lead to selection of clones with a greater ability to repair UV damage than clones which had not undergone UV selection post transfection. RT-PCR analysis was required to investigate which transcripts were being produced by the transfected clones.

Figure 5.4 Correction of UV sensitivity of clones of *Ercc1*-null keratinocytes (Ker.(-/-) following transfection with *Ercc1* MG#11

The UV survival curves for wild type (Ker. +/+), *Ercc1*-deficient (Ker.-/-) murine keratinocytes, three clones which had HAT selection only following transfection with MG#11 are shown in red (clones 7, 3 and 2) and two clones which had HAT and UV selection are shown in blue (clones 5 and 4). Pink, Ker.(-/-); green Ker.(+/+); red, clones from HAT only selection regime; blue, clones from HAT+UV selection regime. Each UV point represents the mean of two separate experiments.


100
10
1
SURVIVAL INDEX
0
5
10
15
20
UV DOSE (Jm-2)
C#7
C#3
C#5
C#2
C#4

5.3.4 Correction of UV hypersensitivity of *Ercc1*-deficient murine keratinocyte pools following transfection with *Ercc1*MG#11

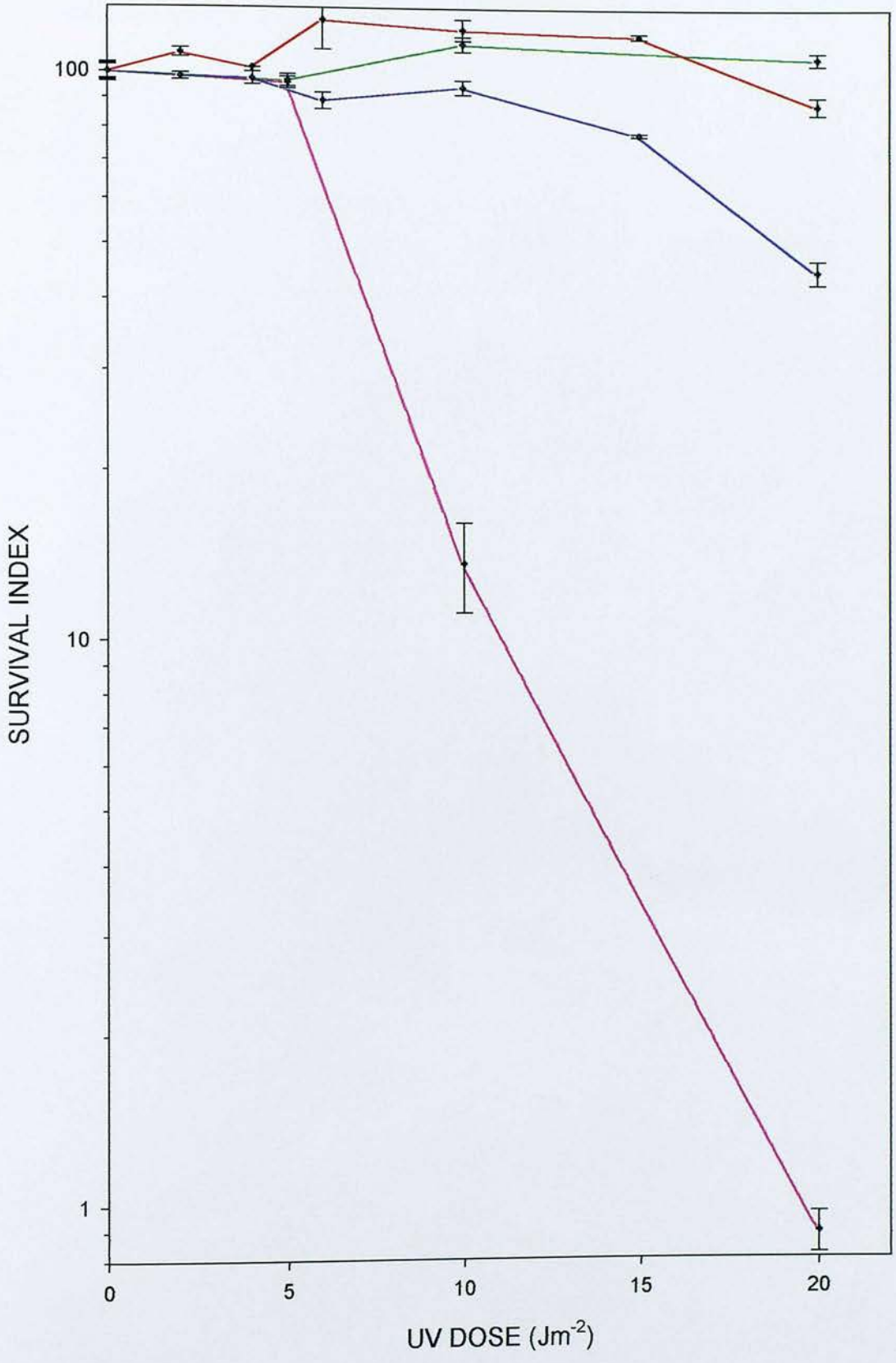
The survival curves of the pools of clones transfected with *Ercc1* MG#11 that underwent HAT or HAT and UV selection post transfection are shown in figure 5.5. The number of colonies from each selection regime that went into each pool are shown in Table 5.1. The survival curves of the two pools show correction of UV hypersensitivity by transfection with *Ercc1* MG#11. The survival index percentages for the pool that underwent HAT only selection following minigene transfection (HAT pool) had the following survival index percentages of 124, 119, 117 and 88% for UV doses of 6, 10, 15 and 20 Jm^{-2}. The survival index percentages for the pool that underwent HAT and UV selection following transfection for the same UV doses are 89, 94, 78 and 45% (D_{50} value = $19Jm^{-2}$) respectively. The survival curve of the pool co-selected with HAT + UV post transfection with MG#11 is lower than the pool selected with HAT. The pool selected with HAT alone has a higher survival index score than the wild type keratinocytes over the middle part of the UV range. The error bars of the pool selected with UV + HAT and the pool selected with HAT overlap at the lower UV doses. At UV dose 10 Jm^{-2} there is not quite a statistically significant difference between the pools (P=0.0515); at UV dose 15 Jm^{-2} there is a very statistically significant difference (P=0.0012) and at UV dose 20 Jm^{-2} the difference between the pools is also very statistically significant. Again, these results are odd because it makes no biological sense that the HAT only selection has resulted in clones with greater correction of UV hypersensitivity than the clones selected with HAT and UV. In pools any sampling effect should be diluted as many clones comprise each pool.

When the survival curves of the pools transfected with *Ercc1*MG#11 are compared to those of the clones (figure 5.6), the curves overlap, indicating very similar UV survival capabilities. When the clones and pool that underwent HAT only selection following transfection are compared the curve of the pool lies within those of the clones, indicating that the UV survival ability of the cells is very similar. The same result is seen when the UV survival curves of the clones and pool selected by HAT and UV post transfection are compared. The survival curve of the pool lies within

those of those clones. It is clear from these results that the clones selected following either selection regime did not have higher than average UV survival capability since the pools of clones have the same UV hypersensitivity correction as the clones themselves. It seems from the results that the presence of the *Ercc1* minigene corrects UV hypersensitivity regardless of post transfection selection regime and that the use of UV selection did not result in clones or pools with greater UV resistance, in fact, the clones and pools selected with HAT only generally had greater correction of UV hypersensitivity. This result makes no biological sense and the possibility of a sampling effect could only be ruled out by repeating the experiment and testing more clones to see how much variation exists. The RT-PCR results would demonstrate which *Ercc1* transcripts were being expressed by the clones and pools investigated. Due to the deletion in the minigene the hypothesis was that these cells would only be expressing the 1.1kb transcript, which made the UV results interesting as the lack of upper transcript in cells which normally have high levels of upper transcript had not impaired the correction of UV hypersensitivity. In fact, when the UV curves of the clones and pools transfected with *Ercc1*MG#11 are compared to those of *Ercc1* MG#13 there is no difference in the effectiveness of the minigenes at correcting UV hypersensitivity, no difference had previously been observed in PF24 or CHO43.3B (TG#1) either.

Figure 5.5 Correction of UV sensitivity of pools of *Ercc1*-null keratinocytes (Ker.(-/-)) following transfection with *Ercc1*MG#11

The UV survival curves for wild type (Ker. +/+), *Ercc1*-deficient (Ker.-/-) murine keratinocytes, one pool which had HAT selection only following transfection with MG#11 (pool HAT) and one pool which had HAT and UV selection (pool HAT + UV) are shown. Each UV point represents the mean of two separate experiments. Pink, Ker.(-/-); green, Ker.(+/+); blue, HAT + UV selected pool; red, HAT selected pool.


100
10
1
SURVIVAL INDEX
0
5
10
15
20
UV DOSE (Jm-2)

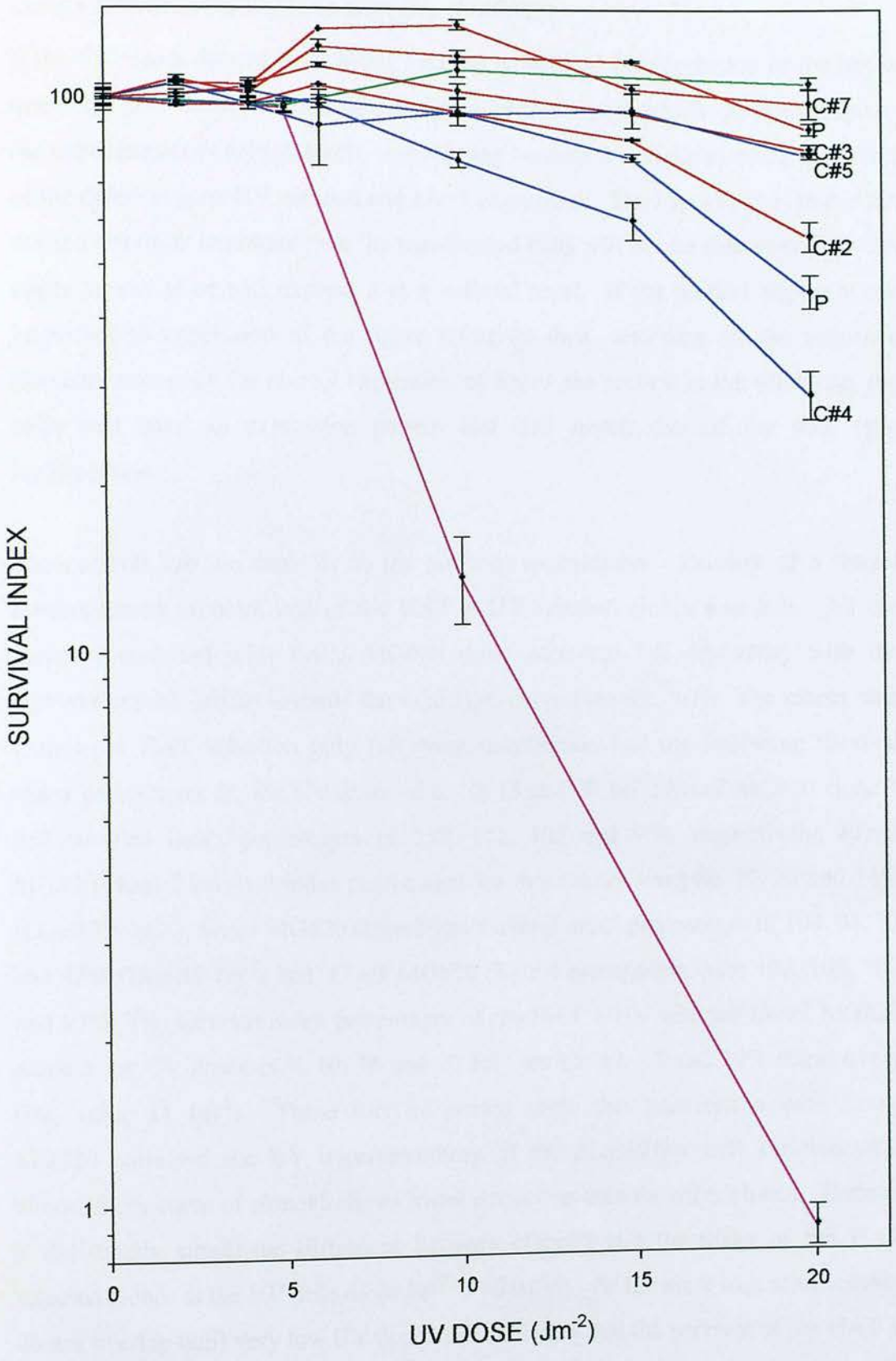

100
10
1
SURVIVAL INDEX
0
5
10
15
20
UV DOSE (Jm-2)
C#7
P
C#3
C#5
C#2
P
C#4

Figure 5.6 Correction of UV sensitivity of clones and pools of *Ercc1*-null keratinocytes (Ker.(-/-)) following transfection with *Ercc1* MG#11

The UV survival curves for wild type (Ker. +/+), *Ercc1*-deficient (Ker. (-/-)) murine keratinocytes, one pool plus three clones (clones, 2, 3 and 7) which had HAT selection only following transfection with MG#13 are shown in red. One pool plus 2 clones (clones 4 and 5) which had HAT and UV selection are shown in blue. Each UV point represents the mean of two separate experiments. Pink, Ker.(-/-); green, ker.(+/+); blue, HAT + UV selected pool and clones; red, HAT selected pool and clones.

5.3.5 Correction of UV hypersensitivity of *Ercc1*-deficient murine keratinocyte clones following transfection with *Ercc1* MG#20

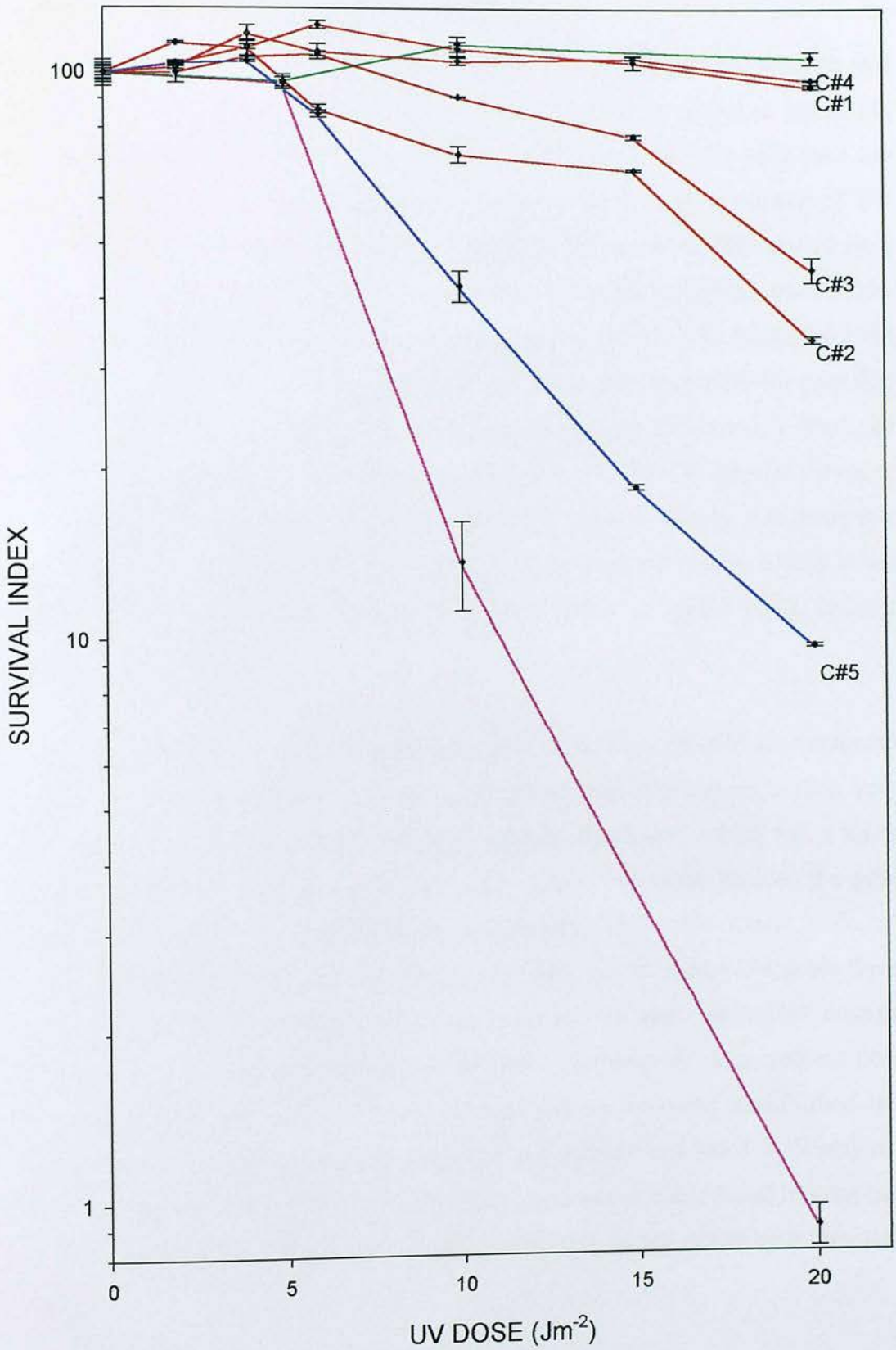
If the CT repeat deleted from *Ercc1* MG#20 is required for expression of the upper transcript then transformation of *Ercc1* null keratinocytes, which normally express the upper transcript at high levels, is interesting because it will demonstrate the effect of the deletion upon UV survival and *Ercc1* expression. The hypothesis is that if the deleted region is important then the transformed cells will not be able to express the upper transcript or will express it at a reduced level. If the deleted region is not important to expression of the upper transcript then, assuming all the sequence elements necessary for correct expression of *Ercc1* are present in the minigene, the cells will have an expression pattern that will match that of the wild type keratinocytes.

The controls are the same as in the previous experiments. Because of a fungal contamination problem one of the HAT + UV selected clones was lost. All the clones transfected with *Ercc1* MG#20 show corrected UV sensitivity with the survival curves shifted towards the wild type curve (see fig. 5.7). The clones that underwent HAT selection only following transfection had the following survival index percentages for the UV doses of 6, 10, 15 and 20 Jm^{-2}; *Ercc1* MG#20 clone 1 had survival index percentages of 112, 111, 105 and 95% respectively, *Ercc1* MG#20 clone 2 survival index percentages for those doses were 86, 72, 67 and 34% (D_{50}=17.5 Jm^{-2}), *Ercc1* MG#20 clone 3 had survival index percentages of 108, 91, 77 and 45% (D_{50}=19 Jm^{-2}) and *Ercc1* MG#20 clone 4 percentages were 109, 105, 107 and 97%. The survival index percentages of the HAT + UV selected *Ercc1* MG#20 clone 5 for UV doses of 6, 10, 15 and 20 Jm^{-2} are 85, 42, 18 and 10% respectively (D_{50} value 11 Jm^{-2}). These survival curves show that transfection with *Ercc1* MG#20 corrected the UV hypersensitivity of the *Ercc1/Hprt* null keratinocytes, although the curve of clone#5 shows lower correction than the other clones. There is a statistically significant difference between clone#5 and the mean of the HAT selected clones at the UV dose of 20 Jm^{-2} (P=0.0186). As the error bars of the clones do not overlap until very low UV doses, this indicates that the survival of the HAT + UV selected clone 5 is very different from the HAT only selected clones, especially

at the higher UV doses. Without information from other clones selected with HAT + UV and transfected with MG#20 it is not possible to tell if this clone had abnormally low expression of the minigene or is a typical survival curve for those conditions. As clone 5 was selected by HAT + UV it is unlikely that the selection regime resulted in selection of a clone with low UV survival. It is possible that the minigene has integrated into a less transcriptionally accessible part of the genome, leading to lower expression and that this statistically significant difference between this clone and the HAT only selected clones is a sampling effect. It makes no biological sense that selection with HAT only should result in clones with higher correction of UV hypersensitivity in contrast to selection with HAT and UV. The clones selected by HAT following transfection with *Ercc1* MG#20 have survival curves that are very similar to that of the wild type keratinocytes in the cases of clones 1 and 4. Clones 2 and 3 also have high UV survival ability. RT-PCR analysis was required to investigate which transcripts the transfected clones were producing.

Figure 5.7 Correction of UV sensitivity of clones of *Ercc1*-null keratinocytes (Ker.(-/-)) following transfection with *Ercc1* MG#20

The UV survival curves for wild type (Ker. +/+), *Ercc1*-deficient (Ker.-/-) murine keratinocytes, four clones which had HAT selection only following transfection with MG#20 are shown in red (clones 1, 2, 3 and 4) and one clone which had HAT and UV selection is shown in blue (clones 5). Pink, Ker.(-/-); green Ker.(+/+; red, clones from HAT only selection regime; blue, clone from HAT+UV selection regime. Each UV point represents the mean of two separate experiments.


100
10
1
SURVIVAL INDEX
0
5
10
15
20
UV DOSE (Jm-2)
C#4
C#1
C#3
C#2
C#5

5.3.6 Correction of UV hypersensitivity of *Ercc1*-deficient murine keratinocyte pools following transfection with *Ercc1*MG#20

The survival curves of the pools of clones transfected with *Ercc1* MG#20 that underwent HAT or HAT and UV selection post transfection are shown in figure 5.8. The number of colonies from each selection regime that went into each pool are shown in Table 5.1. The survival curves of the two pools show correction of UV hypersensitivity by transfection with *Ercc1* MG#20. The survival index percentages for the pool that underwent HAT only selection following minigene transfection (HAT pool) had the following survival index percentages of 99, 100, 88 and 56% for UV doses of 6, 10, 15 and 20 Jm^{-2}. The survival index percentages for the pool that underwent HAT and UV selection following transfection for the same UV doses are 93, 89, 86 and 62% respectively. There is overlap between the UV survival curves of the two pools and both show equal correction of UV hypersensitivity, indicating that including UV in the selection regime post transfection did not lead to a shift in the population of cells towards those with greater ability to repair DNA damage following UV irradiation.

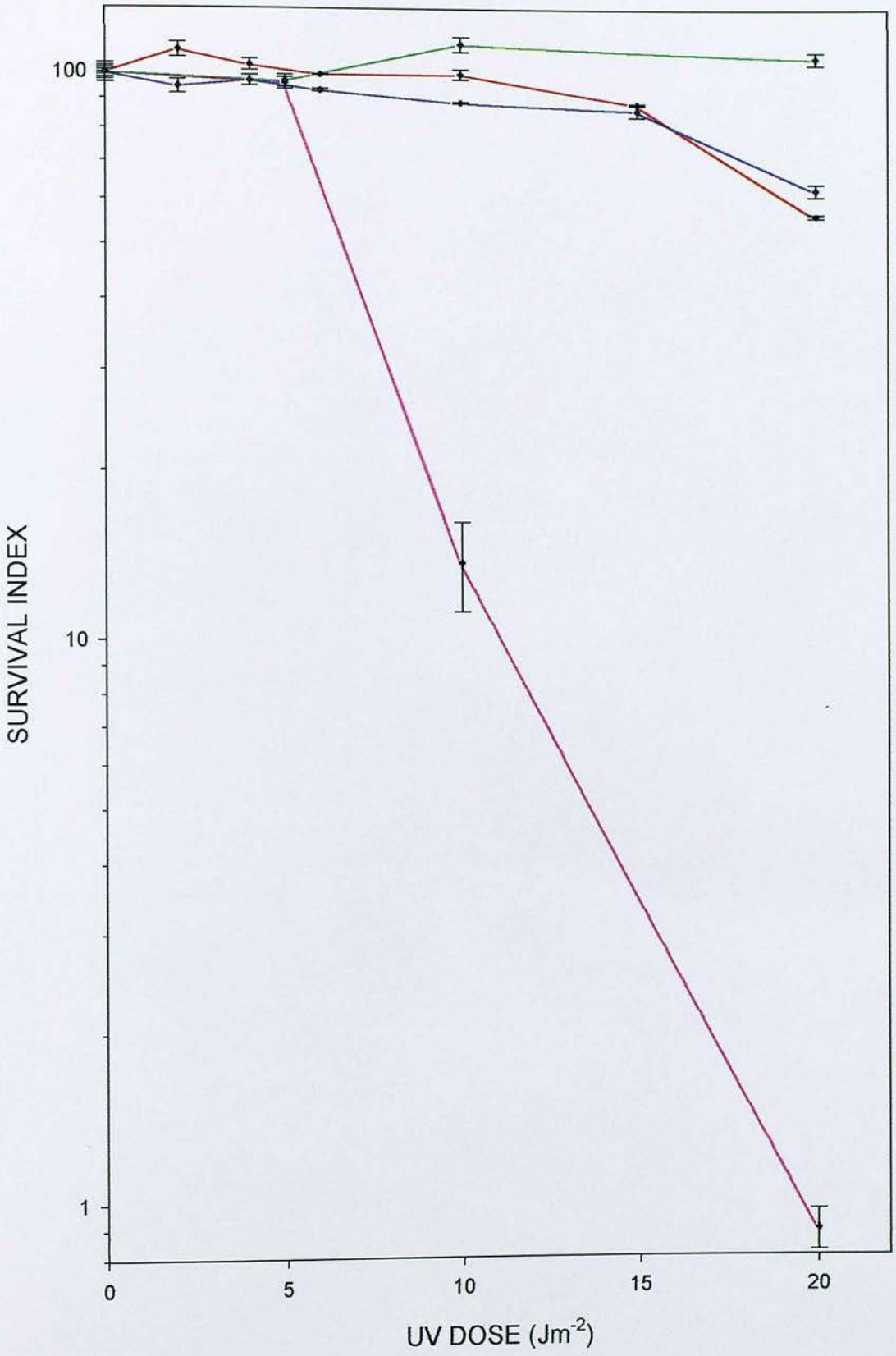
When the survival curves of the pools transfected with *Ercc1*MG#20 are compared to those of the clones (figure 5.9), the curve distribution overlaps, indicating very similar UV survival capabilities, with the exception of clone#5 which has a lower UV survival curve than the other pools and clones. The error bars of the pool selected with HAT + UV and the clone selected with HAT + UV (clone 5) do not overlap above the low UV dose of 5 Jm^{-2}, indicating that at higher UV doses there was a difference in the survival between the clone and the pool. At the UV dose of 20 Jm^{-2} there is an extremely significant difference between the clone and the pool ($P= 0.0009$). However, if may be that clone#5 was not correctly transformed and that other clones transfected with the minigene and selected with HAT +UV may not have such low correction of UV hypersensitivity. It is not possible to tell how typical clone#5 is of cells transfected and selected in this way as this could be a sampling effect.

When the clones and pool that underwent HAT only selection following transfection are compared the survival curve of the pool lies within those of the clones, indicating that the UV survival ability of the cells is very similar. The fact that the clone selected by HAT + UV has the lowest UV survival of all the clones and pools (including those selected by HAT alone) is odd as the clone obviously survived the UV selection regime post electroporation. It is possible that the minigene inserted into a site that was transcriptionally difficult to access yet the later results of the RT-PCR (figure 5.12) show that the clone is expressing *Ercc1* transcripts at levels comparable to the other clones.

The distribution of survival curves indicate that selection by UV does not influence the UV survival of the resulting clones when compared to clones selected by HAT (which due to the original 3:1 *Ercc1: Hprt* DNA ratio are likely to contain the *Ercc1* gene if they survive HAT selection). The RT-PCR results would demonstrate which *Ercc1* transcripts were being expressed. Due to the deletion in the minigene the hypothesis was that these cells would be expressing only the 1.1kb transcript or both transcripts with a reduced level of upper transcript when compared to the cells transformed with *Ercc1* MG#13.

Figure 5.8 Correction of UV sensitivity of pools of *Ercc1*-null keratinocytes (Ker.(-/-)) following transfection with *Ercc1* MG#20

The UV survival curves for wild type (Ker. +/+), *Ercc1*-deficient (Ker.-/-) murine keratinocytes, the pool which had HAT selection only following transfection with MG#20 (pool HAT) and the pool which had HAT and UV selection (pool HAT + UV) are shown. Each UV point represents the mean of two separate experiments. Pink, Ker.(-/-); green, Ker.(+/+); blue, HAT + UV selected pool; red, HAT selected pool.


100
10
1
SURVIVAL INDEX
0
5
10
15
20
UV DOSE (Jm-2)

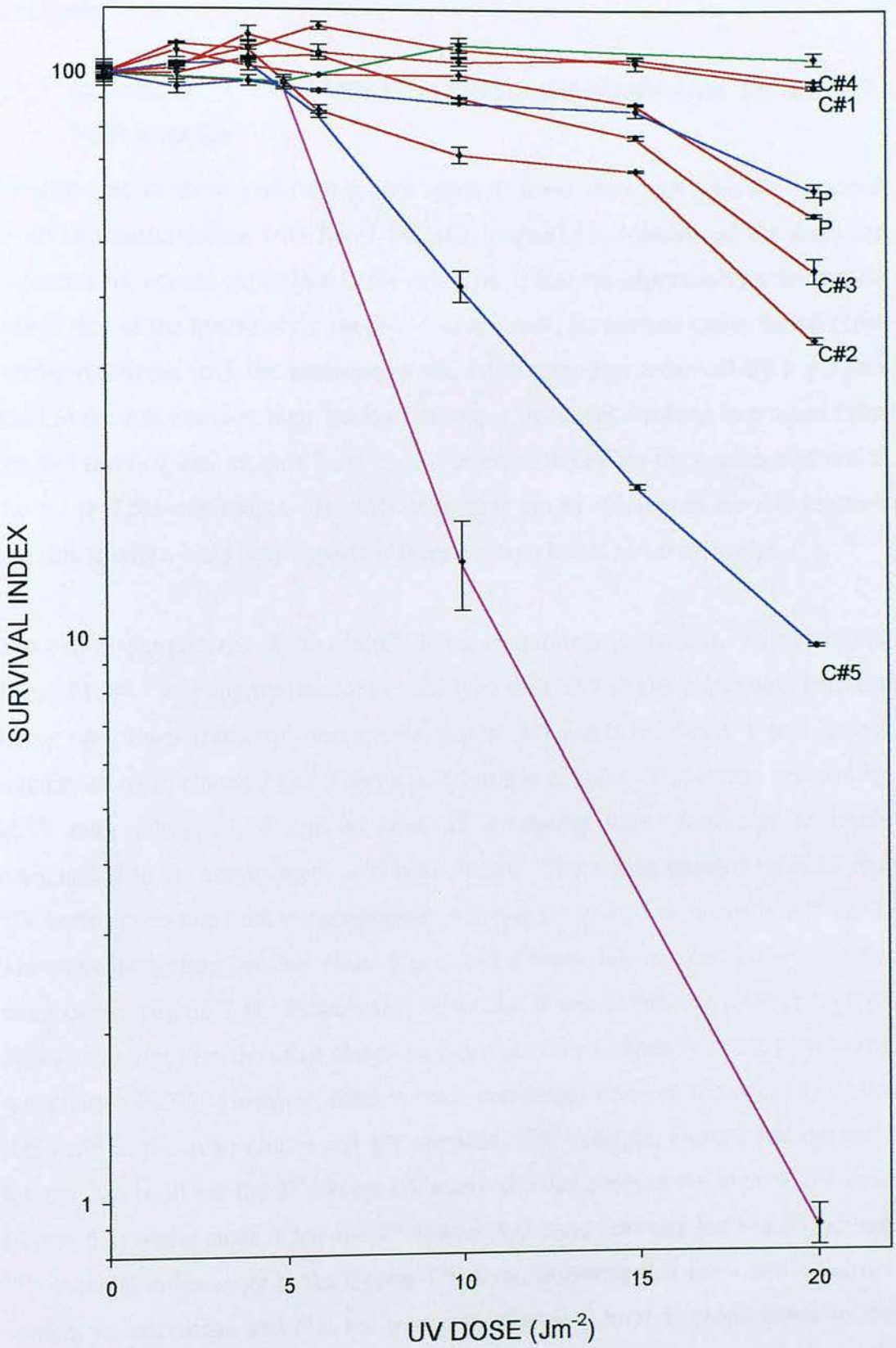

100
10
1
SURVIVAL INDEX
0
5
10
15
20
UV DOSE (Jm-2)
C#4
C#1
P
P
C#3
C#2
C#5

Figure 5.9 Correction of UV sensitivity of clones and pools of *Ercc1*-null keratinocytes (Ker.(-/-)) following transfection with *Ercc1* MG#20

The UV survival curves for wild type (Ker. (+/+)), *Ercc1*-deficient (Ker.(-/-)) murine keratinocytes, one pool plus four clones (clones 1, 2, 3 and 4) which had HAT selection only following transfection with MG#20 are shown in red. One pool plus 1 clone (clone 5) which had HAT and UV selection are shown in blue. Each UV point represents the mean of two separate experiments. Pink, Ker.(-/-); green, Ker.(+/+); blue, HAT + UV selected pool and clones; red, HAT selected pool and clones.

5.4 Expression of minigenes in Ercc1-deficient keratinocytes – RT-PCR analysis

5.4.1 Expression of *Ercc1* MG#13 in transfected keratinocyte clones– RT-PCR analysis

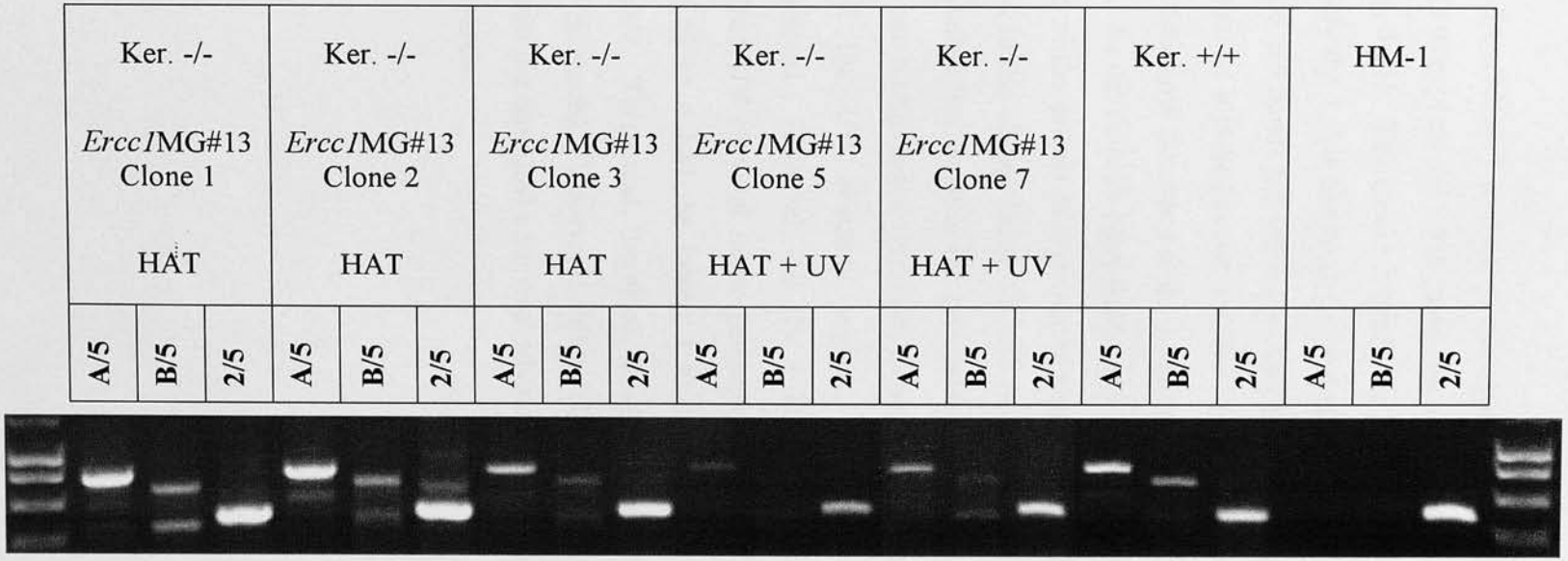
As wild type keratinocytes have a high upper to lower transcript ratio the expected result of transformation with *Ercc1* MG#13, assuming it contains all the elements necessary for correct expression in the cell type, is that the expression pattern would mimic that of the keratinocyte control. There should be product bands for all three primer reactions, with the presence of the 1.5kb transcript indicated by a product band in the A/5 reaction, high levels of the upper transcript resulting in product from the B/5 reaction and product from the 2/5 reaction indicating the presence of either the 1.1 or 1.5kb transcripts. The B/5 reaction is not as efficient as the A/5 reaction and this is why a band only appears if there are high levels of 1.5 transcript.

The expression patterns of the controls in the experiment are correct. Expression of *Ercc1* MG#13 was appropriate for the cell type with all 5 clones expressing both the upper and lower transcript, there were strong A/5 bands in clones 1 and 2, less intense bands in clones 3 and 7 and a pale band in clone 5. The clones selected by HAT only (clones 1, 2 and 4) were all producing upper transcript at levels comparable to the keratinocyte wild type control. The clones selected by HAT and UV were expressing both transcripts although clone 5 had a low intensity A/5 band. This was interesting because clone 5 also had a lower UV survival curve than the other clones (figure 5.1). Perhaps this is because it was expressing a lower level of upper transcript than the other clones (it is not possible to know without performing quantitative PCR). However, there is not a correlation between the intensity of the A/5 band in the other clones and UV survival. For example, clone 1 has the most intense A/5 band but the 2nd lowest UV survival index score at the highest UV dose (figure 5.1) while clone 7 has the 2nd lowest A/5 band intensity but the 3rd highest UV survival index score at the highest UV dose, assuming that the 1.5kb transcript confers an advantage and that the transcript expressin level is proportional to the intensity of the product band. Because the RT-PCR used was not quantitative, the

intensity of the product bands may not correlate with the expression level of the corresponding transcript. The band intensity pattern is the opposite of what would have been expected if the presence of the upper transcript confers an advantage when repairing UV damage. If UV irradiation was selecting clones on that basis then the hypothesis would be that the clones which were co-selected with HAT and UV would have more intense A/5 bands than those selected by HAT alone. However, if the 1.1kb is the most important for DNA damage repair capability then the UV irradiation would select clones with a lower 1.5kb to 1.1kb transcript ratio and the RT-PCR pattern would be correct for that hypothesis.

Figure 5.10 Expression of *Ercc1* MG#13 in keratinocyte clones - RT-PCR

RT-PCR analysis of RNA extracted from *Ercc1*-null keratinocyte clones transfected with *Ercc1*MG#13. Total RNA was prepared and reverse transcribed using oligo-dT primers. The *Ercc1* cDNA was then PCR amplified from the cDNA pool of each clone. Three PCR reactions were performed on cDNA from each clone using the following forward primers respectively; primer 'A' is specific a region upstream of the normal transcriptional start site, primer 'B' is specific to a region upstream of the *Kpn1* restriction site and primer '2' is specific to exon 2. Primer '5' is specific to a region in exon 5 and is the reverse primer for each of the three reactions.


Ker. -/-
*Ercc1*MG#13 Clone 1
HAT
Ker. -/-
*Ercc1*MG#13 Clone 2
HAT
Ker. -/-
*Ercc1*MG#13 Clone 3
HAT
Ker. -/-
*Ercc1*MG#13 Clone 5
HAT + UV
Ker. -/-
*Ercc1*MG#13 Clone 7
HAT + UV
Ker. +/+
HM-1
A/5
B/5
2/5

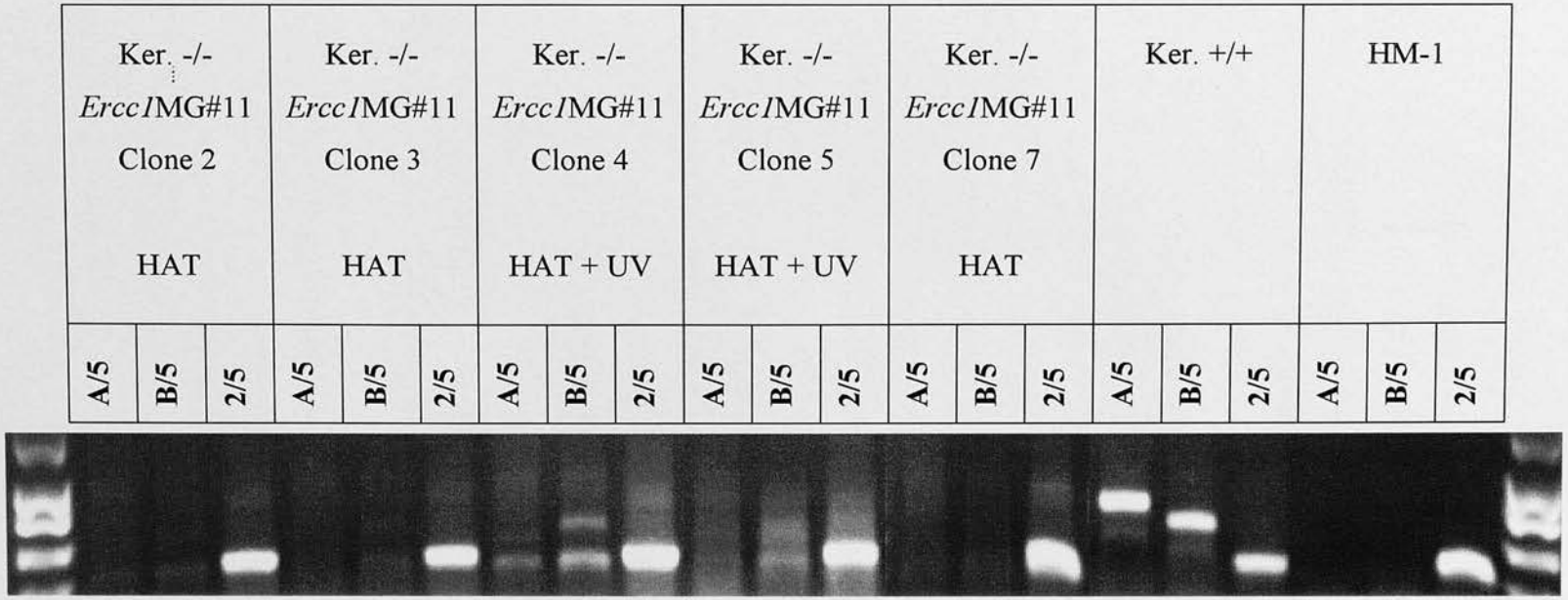
5.4.2 Expression of *Ercc1* MG#11 in transfected keratinocyte clones– RT-PCR analysis

The deletion in *Ercc1*MG#11 resulted in the expected expression pattern with none of the clones able to express the 1.5kb transcript and, therefore, lacking A/5 bands in the RT-PCR (fig. 5.11). The clones selected by HAT alone had only the 2/5 band, indicating that only the 1.1kb transcript was being produced. The clones selected by HAT and UV had 2/5 bands but there was also a B/5 band in both of these clones. This was intriguing as it indicates the possible presence of a transcript that initiated above the 1.1kb transcript but was not the 1.5kb transcript as the skin TSS is missing in this minigene. As the controls have the correct pattern it is unlikely that this is due to contamination of the primer master mix or an artefact of the reaction. B/5 bands had not appeared in the earlier *Ercc1* MG#13 transfections of fibroblasts and CHO cells. It was possible that because the minigene was being expressed in a cell line which normally has a high upper transcript expression and because the clones were the products of the UV selection regime, perhaps abnormal clones were preferentially selected. However, the UV + HAT selected clones did not have significantly higher UV survival index percentages than the clones selected with HAT alone and clone 4 had the lowest UV survival curve although it had the brightest B/5 bands. There was, therefore, no correlation between the presence of B/5 product and increased UV survival ability. This suggests that the cells with the B/5 products would not have had a survival advantage under the UV selection regime following transfection.

Figure 5.11 Expression of *Ercc1* MG#11 in keratinocyte clones - RT-PCR

RT-PCR analysis of RNA extracted from *Ercc1*-null keratinocyte clones transfected with *Ercc1*MG#11. Total RNA was prepared and reverse transcribed using oligo-dT primers. The *Ercc1* cDNA was then PCR amplified from the cDNA pool of each clone. Three PCR reactions were performed on cDNA from each clone using the following forward primers respectively; primer 'A' is specific a region upstream of the normal transcriptional start site, primer 'B' is specific to a region upstream of the *Kpn1* restriction site and primer '2' is specific to exon 2. Primer '5' is specific to a region in exon 5 and is the reverse primer for each of the three reactions.

The lower band in the A/5 and B/5 lanes of clone 4 are likely to be spills from the 2/5 reaction of either clone 3 or clone 5.


Ker. -/-
*Ercc1*MG#11
Clone 2
HAT
Ker. -/-
*Ercc1*MG#11
Clone 3
HAT
Ker. -/-
*Ercc1*MG#11
Clone 4
HAT + UV
Ker. -/-
*Ercc1*MG#11
Clone 5
HAT + UV
Ker. -/-
*Ercc1*MG#11
Clone 7
HAT
Ker. +/+
HM-1
A/5
B/5
2/5

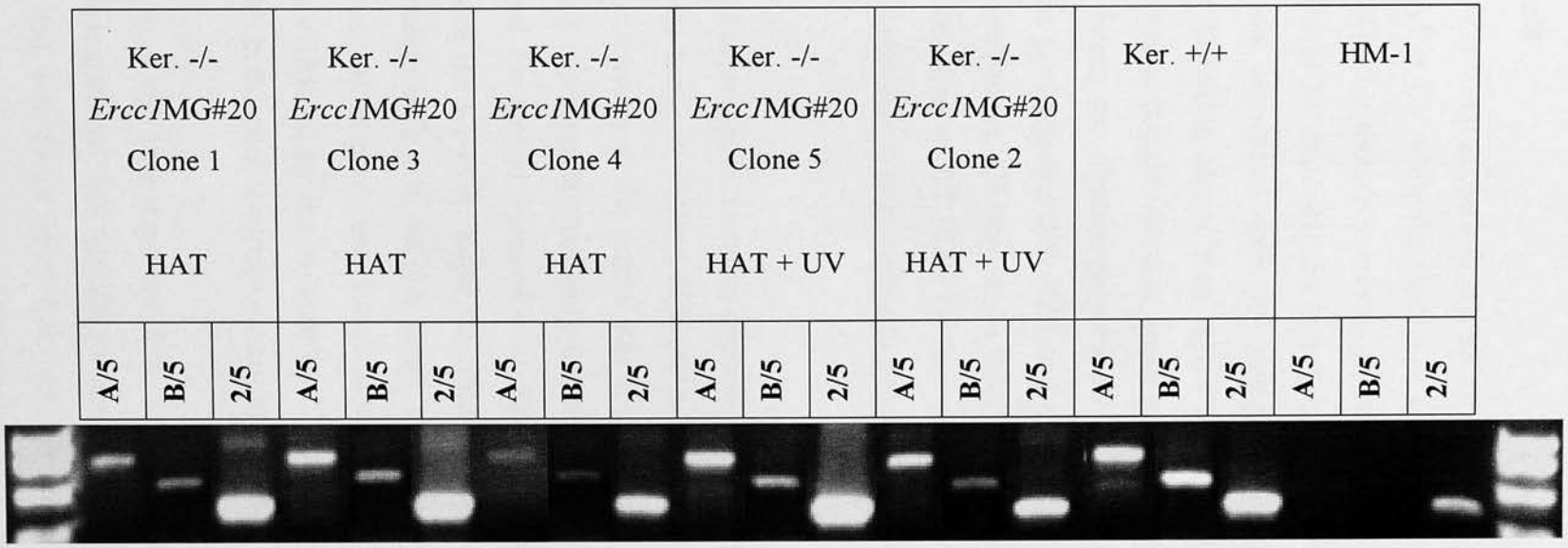
5.4.3 Expression of *Ercc1* MG#20 in transfected keratinocyte clones– RT-PCR analysis

If the CT repeat region deleted in *Ercc1* MG#20 was required for expression of the upper transcript then keratinocytes transformed with this minigene should not express the upper transcript or express it at a reduced level in comparison with those transformed with *Ercc1* MG#13 (the expression level could only be verified by means of quantitative RT-PCR). When the RT-PCR results are examined (fig. 5.12) it is apparent that the deletion has not prevented expression of the upper transcript in transformed keratinocytes. The three clones selected by HAT and the two clones selected by HAT + UV following transfection all have A/5 and B/5 bands indicating that they are producing the upper transcript. There are also B/5 bands, which suggests that the upper transcript is probably being expressed at high levels. Remembering that the observations upon band intensity is presented for interest only and in the ansence of data from quantitative RT-PCR these observations must not be over-interpreted, the clones selected by HAT and UV have high intensity A/5 bands and two of the three clones selected by HAT and UV (clones 1& 3) have high intensity A/5 bands while one has a slightly less intense A/5 band. These observations may suggest that the lack of the CT repeat did not impair the ability of the transformed cells to produce upper transcript as would have been expected had this region been a crucial part of the skin specific transcript promoter but the actual expression level of the transcripts has not been ascertained so this is speculation only. The results do demonstrate that it is possible to express the skin-specific transcript despite deletion of the CT repeat region.

There is no correlation between the intensity of the A/5 band and the UV survival ability of the clone. The clones in decreasing order of A/5 band intensity are 5, 3, 2, 1 and 4. In order of highest UV survival index score at the highest UV dose, the clones in decreasing order are 4, 1, 3, 2 and 5. Those with the highest A/5 band intensity had the lowest UV survival.

Figure 5.12 Expression of *Ercc1* MG#20 in keratinocyte clones - RT-PCR

RT-PCR analysis of RNA extracted from *Ercc1*-null keratinocyte clones transfected with *Ercc1*MG#20. Total RNA was prepared and reverse transcribed using oligo-dT primers. The *Ercc1* cDNA was then PCR amplified from the cDNA pool of each clone. Three PCR reactions were performed on cDNA from each clone using the following forward primers respectively; primer 'A' is specific a region upstream of the normal transcriptional start site, primer 'B' is specific to a region upstream of the *Kpn1* restriction site and primer '2' is specific to exon 2. Primer '5' is specific to a region in exon 5 and is the reverse primer for each of the three reactions.


Ker. -/- *Ercc1*MG#20 Clone 1 HAT
Ker. -/- *Ercc1*MG#20 Clone 3 HAT
Ker. -/- *Ercc1*MG#20 Clone 4 HAT
Ker. -/- *Ercc1*MG#20 Clone 5 HAT + UV
Ker. -/- *Ercc1*MG#20 Clone 2 HAT + UV
Ker. +/+
HM-1
A/5
B/5
2/5

5.5 Expression of Ercc1 minigenes in transfected keratinocyte pools- RT-PCR analysis

There were six pools – two for each minigene. In each minigene pair, one pool was selected with HAT + UV while the other was selected with HAT only (fig.5.12). Because *Ercc1* MG#13 is thought to contain all the sequence required for expression of both transcripts in keratinocytes the hypothesis is that both transcripts will be expressed in pools transfected with this minigene. If the UV selection post transfection was favouring clones with higher DNA repair ability and this was because they expressed a higher level of upper transcript then, assuming that more intense product bands are directly proportional to transcript expression level, a comparison of the pool selected with HAT and the pool selected with HAT + UV would show a more intense A/5 band in the HAT + UV selected pool. However, in the absence of quantitative RT-PCR the following observations based upon the premise that band intensity is proportional to transcript expression level remain mere speculation.

Minigene#13 was expressed in a manner appropriate for the transfected cells in both pools. There were bands in all three PCR reactions for both pools and these bands were of a similar intensity to the control Ker +/+ reactions. This suggests that the minigene is being expressed appropriately for the cell type and contains all of the elements required for correct expression of the 1.5kb transcript. There was no difference between the RT-PCR results for either pool selection method indicating that the co-selection with UV did not result in a pool with a different *Ercc1* expression pattern. Both pools transfected with *Ercc1* MG#13 had similar DNA repair ability as evidenced by the overlapping UV survival curves and both pools expressed the MG in the same transcript pattern and levels.

Ercc1 MG#11 lacks the skin transcript TSS. This means that it should not be possible for cells transfected with this MG to express the 1.5kb transcript. The pools of cells transfected with *Ercc1* MG#11 do not have a product band for the A/5 reaction indicating that the upper transcript is not being expressed. Both pools have 2/5 bands indicating the presence of the normal 1.1kb transcript. However, there is a

pale band in the B/5 reaction of the HAT+UV selected pool and a very pale B/5 band is also present in the HAT selected pool. This is intriguing as it is a possible indication of a transcript which is initiating between the primer A and primer B hybridisation sites and, therefore, upstream of the normal transcript. This cannot be the 1.5kb transcript as the TSS of that transcript is not present in *Ercc1* MG#11. There is no B/5 band in the HM-1 control indicating that the PCR mix was not contaminated. If the production of the transcript detected by the B/5 reaction confers an advantage under conditions of UV exposure then perhaps the UV selection regime caused cells expressing the MG in such a way to be favoured over cells expressing only the 1.1kb transcript (the B/5 band is brighter in the HAT + UV pool when compared to the HAT pool, although without quantitative RT-PCR this must not be over-interpreted). Any such selective advantage is not apparent in the UV survival assay of the two pools (fig. 5.5) as the pool that was selected with HAT had a higher UV survival curve that the pool selected with HAT and UV. The presence of the B/5 band in HAT + UV selected clones was discussed in section 5.4.2 and the same explanation applies here. The band either indicates that a transcript is initiating upstream of the normal transcript and this could be due to selective pressure for such a transcript in the absence of the normal 1.5kb transcript in keratinocytes. Alternatively, the transcript could be initiating in the plasmid sequence upstream of the primer B hybridisation site, in which case it is merely an artefact.

Ercc1 MG#20 lacks the CT repeat region that could be part of the promoter of the 1.5kb transcript. If this region is required for correct expression of the transcript then RT-PCR on the pools transfected with this MG should show reduced expression or complete absence of the 1.5kb transcript evidenced by the level of the A/5 product band. If the UV selection regime following transfection was responsible for the selection of cells which had higher UV survival ability and this was because of the level of upper transcript then this would also be detected in the RT-PCR when the HAT selected pool results and HAT + UV pool results were compared.

Ercc1 MG#20 was expressed appropriately for the cell type with A/5 product bands in both pools. Again, observations upon the band intensity remain speculations when

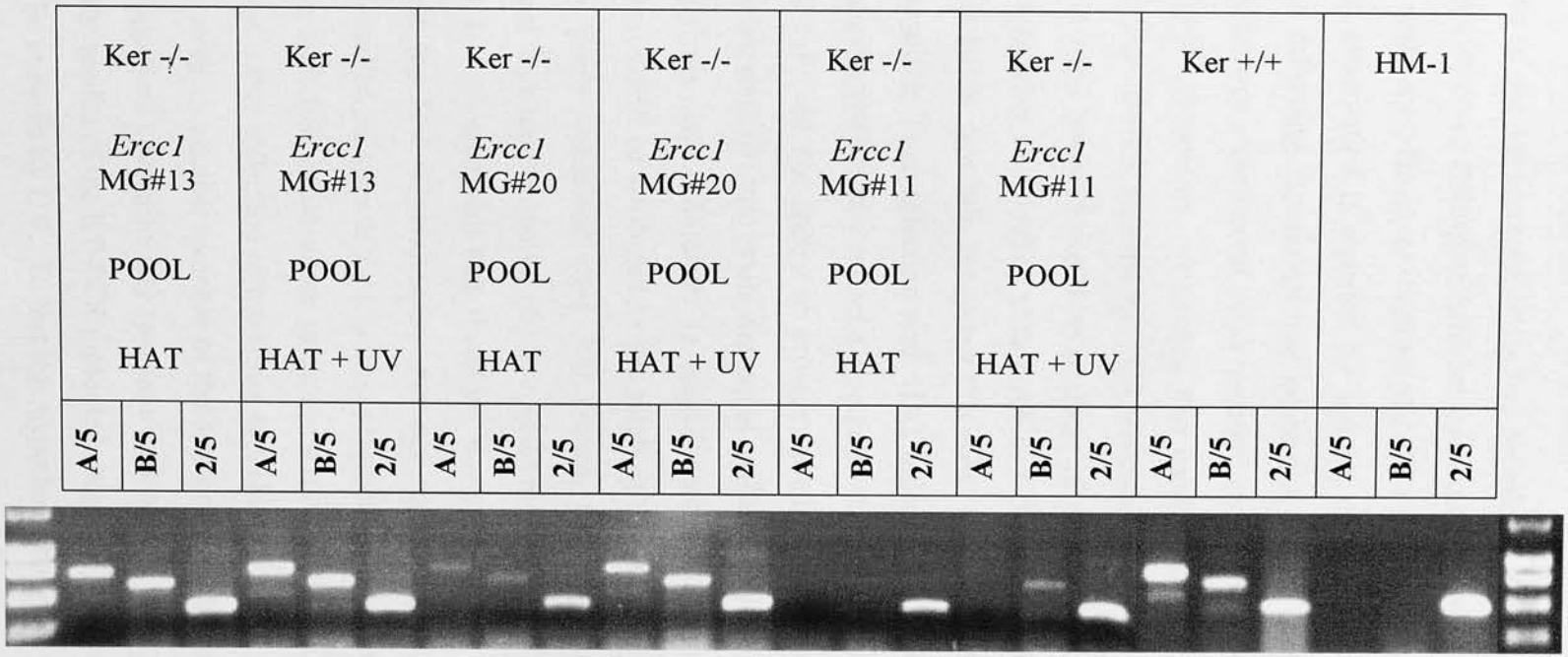
correlated with transcript expression levels in the absence of quantitative RT-PCR. The intensity of the A/5 bands in the HAT + UV pool was similar to that of the Ker. (+/+) control but the intensity of the A/5 band in the HAT selected pool was lower than in the Ker. (+/+) control and the HAT + UV selected pool. The intensity of the A/5 bands in both pools transfected with *Ercc1* MG#20 were lower than the pools transfected with *Ercc1* MG#13 and this could indicate reduced 1.5kb transcript production due to the absence of the CT repeat region in *Ercc1* MG#20 (without follow-up experiments using quantitative RT-PRC this remains speculation). However, because there is a difference between the results in the RT-PCR patterns for the two pools transfected with *Ercc1* MG#20 it is also possible that the UV+HAT selection regime is biased towards cells expressing the upper transcript. Because the UV survival curves of the two pools were so similar (fig. 5.8) any survival advantage under UV conditions conferred by the presence of upper transcript is not apparent. It is, therefore, possible that deletion of the CT repeat region could cause a reduction in upper transcript production but does not preclude production of this transcript entirely.

The following observations on band intensity are reported for interest only and, in the absence of quantitative RT-PCR data no conclusions may be drawn. The A/5 bands in the *Ercc1* MG#13 pools are equally bright for both selection regimes. However, it is interesting that the pools which underwent HAT + UV selection had more intense A/5 bands than the HAT only selected pools in the case of cells transformed with MG#20, indicating that the UV+HAT pools could potentially be expressing the upper transcript at higher levels if the band intensity is proportional to transcript expression level. If the upper transcript increases the DNA repair ability of the cells then this was not apparent in the UV survival assays performed upon the pools. In order to investigate the possibility that UV could induce upper transcript production the pool set with the greatest difference in upper transcript band intensity, the *Ercc1* MG#20 pools, were used in a follow-up experiment (5.6). The pools should be more representative of cells transformed with the minigenes than the clones alone, so these results suggest that *Ercc1* MG#13 contains all the sequence elements required for correct expression of the upper transcript in keratinocytes and

correct the UV hypersensitivity of transformed cells. The deletion in *Ercc1* MG#20 may cause slightly reduced upper transcript production (if band intensity is proportional to transcript expression level) in transformed keratinocytes but this minigene is capable of correcting the UV hypersensitivity of transformed cells equally effectively as transformation with *Ercc1* MG#13. *Ercc1* MG#11 lacks the skin TSS so keratinocytes transformed with this MG cannot express the upper transcript. However, transfection with this minigene was equally effective at correcting the UV hypersensitivity of the cells as transfection with *Ercc1* MG#13 and 20. This was surprising since keratinocytes normally express the upper transcript at high levels. The absence of this transcript did not reduce the UV survival levels of these cells.

Figure 5.13 Expression of *Ercc1* minigenes in keratinocyte pools - RT-PCR

RT-PCR analysis of RNA extracted from *Ercc1*-null keratinocyte pools transfected with *Ercc1* minigenes #11, #13 and #20. Total RNA was prepared and reverse transcribed using oligo-dT primers. The *Ercc1* cDNA was then PCR amplified from the cDNA of each pool. Three PCR reactions were performed on cDNA from each pool using the following forward primers respectively; primer 'A' is specific a region upstream of the normal transcriptional start site, primer 'B' is specific to a region upstream of the *Kpn1* restriction site and primer '2' is specific to exon 2. Primer '5' is specific to a region in exon 5 and is the reverse primer for each of the three reactions.


Ker -/-
Ercc1 MG#13
POOL
HAT
Ker -/-
Ercc1 MG#13
POOL
HAT + UV
Ker -/-
Ercc1 MG#20
POOL
HAT
Ker -/-
Ercc1 MG#20
POOL
HAT + UV
Ker -/-
Ercc1 MG#11
POOL
HAT
Ker -/-
Ercc1 MG#11
POOL
HAT + UV
Ker +/+
HM-1
A/5
B/5
2/5

5.6 Possible UV induction of upper transcript in keratinocyte pools transfected with Ercc1 minigenes – RT-PCR

RT-PCR analysis of the keratinocyte pools had revealed a difference between the expression pattern observed following different selection regimes in the skin-specific transcript band intensity. However, because the RT-PCR technique used was not quantitative, this observation is reported for interest only and must not be over-interpreted. The following experiment was interesting but no conclusions may be drawn without follow-up experiments using quantitative RT-PCR. This is explained in more detail in the discussion. Assuming that band intensity is proportional to transcript level (remembering that the RT-PCR technique is not qualitative and that the differences in band intensity could be artefacts of the RT-PCR), the pools that were produced following co-selection with HAT and UV had brighter skin-specfic transcript product bands, possibly indicating that these pools expressed more 1.5kb than the pools resulting from selection with HAT alone. The higher skin-specific product band intensity (which may reflect a greater production of upper transcript) in the UV selected cells did not appear to provide an advantage in the UV survival assays as the UV survival of these pools was equal to that of the HAT selected pools. In the clonal analysis such differences in band intensity (and possibly transcript expression pattern) could be attributed to the small sample number but in the pool analysis all the pools contained over 500 clones, meaning that the resultant expression pattern was representative of the transformed cells. The difference in expression in the UV selected pool was, therefore, worthy of further investigation. It was possible that the UV survival assay was not replicating the environmental conditions that resulted in selection of clones expressing upper transcript at high levels over those with low or no upper transcript expression (again, assuming the band intensity was a true reflection of transcript expression levels). At this stage of the experiments, work by another member of the Melton Lab. had suggested that the upper transcript appeared approximately two days post-partum in the mouse skin and this, along with the results of the RT-PCR presented, suggested that production of the transcript could be induced by UV. To test this hypothesis the pool transfected with *Ercc1* MG#20 was chosen for further study. This was because there the expression patterns from the two selection regimes resulted in the greatest difference in skin-

specific transcript product band intensity, possibly reflecting greater upper transcript expression in this sample. The hypothesis was that if the HAT selected pool was irradiated with UV then the RT-PCR band intensity pattern would change to that of the UV + HAT selected pool. If the presence of the upper transcript conferred a survival advantage upon the cells under conditions of UV exposure and the level of transcript was directly proportional to band intensity then UV irradiation would either result in expression of the upper transcript, should induction of expression be possible, or selection for cells expressing higher levels of upper transcript in the pool.

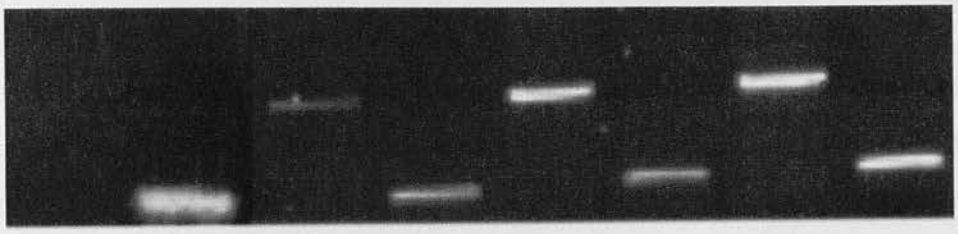
The HAT selected pool was expanded and split into two immediately prior to the experiment. One half received multiple UV-C doses of $6.5Jm^{-2}$ while the other half had no UV exposure. RT-PCR 24 hours following the final UV exposure demonstrated that the pool that had been UV irradiated had a 4 fold higher skin-specific product band intensity than the pool that had not been UV irradiated. If the level of product is a true reflection of the level of transcript then this is an indication of UV induction but correct analysis of the results from this non-quantitative RT-PCR is in terms of absence or presence of product only and in that case there is no difference. The suggested increase in expression of the skin-specific transcript can only be proven or refuted by means of follow-up quantitative RT-PCR. If band intensity was reflecting transcript expression level then the UV irradiated cells had more upper transcript than the non-irradiated cells and UV irradiation appears to either be capable of inducing upper transcript expression or selecting for cells that express the upper transcript at higher levels. If this is the case then the presence of the upper transcript in skin could be explained by the biological requirement of the skin for DNA damage repairing following UV irradiation from sunlight. However, these observations on band intensity changes following UV exposure are reported for interest only and in the absence of quantitative RT-PCR no conclusions may be drawn from this experiment.

Figure 5.14 Possible induction of *Ercc1* upper transcript expression by means of UV irradiation.

A HAT selected pool transfected with *Ercc1* MG#20 (section 5.5) was split into two halves. One half received multiple UV-C doses of $6Jm^2$. 24 hours following the final UV irradiation RNA samples were extracted from both the UV irradiated pool and the non UV-irradiated control for RT-PCR analysis. The 2/5 reaction gives a positive result for both the 1.1 and 1.5kb transcripts. The A/5 reaction gives a positive result for only the skin-specific 1.5kb transcript. Two controls were included; Ker.(+/+) expresses both transcripts and HM-1 (ES cell) express the 1.1kb transcript.

Total RNA was prepared and reverse transcribed using oligo-dT primers. The *Ercc1* cDNA was then PCR amplified from the cDNA pool. Three PCR reactions were performed on cDNA from each clone using the following forward primers respectively; primer 'A' is specific a region upstream of the normal transcriptional start site, primer 'B' is specific to a region upstream of the *Kpn1* restriction site and primer '2' is specific to exon 2. Primer '5' is specific to a region in exon 5 and is the reverse primer for each of the three reactions.

HM-1 (ES cells)		Ker -/- POOL *Ercc1* MG#20 NO UV		Ker -/- POOL *Ercc1* MG#20 UV		Ker +/+	
A/5	**2/5**	**A/5**	**2/5**	**A/5**	**2/5**	**A/5**	**2/5**



5.7 Conclusions

Transfection with *Ercc1* MG#11 may not have corrected the UV hypersensitivity of *Ercc1*-deficient keratinocyte cells as effectively as transfection with *Ercc1* MG#13 and #20, as fewer cells survived the UV selection post electroporation in this group. This is demonstrated by the low ratio of HAT+UV selected colonies to HAT selected colonies in the *Ercc1* MG#11 pool compared to the *Ercc1* MG#13 and #20 pools. The ratio of HAT + UV selected colonies to HAT selected colonies in the pools of cells transfected with minigenes #13 and #20 was almost double that of the pool transfected with *Ercc1* MG#11. This difference may be indicative of the importance of the upper transcript to keratinocyte survival of the UV selection regime post electroporation, since the deletion in minigene#11 results in an inability to express the 1.5kb transcript, the primary transcript in wild type keratinocytes. This large ratio difference (almost 2-fold) between minigenes was not observed in the other cell types studied (in which the primary transcript is the 1.1kb, normal transcript rather than the 1.5kb skin-specific transcript). However, without repeating the experiments with a greater number of dishes plated with cells for each transfection, no conclusions regarding the statistical significance of the ratio differences can be drawn.

Unlike the results of the UV survival assays performed in the PF24 and CHO43.3B (TG#1) experiments, transfection of *Ercc1*-deficient keratinocytes with the minigenes did not result in as much clustering of UV survival curves around the wild-type curve. This could possibly indicate that the minigenes were not as effectively correcting the UV hypersensitivity of the cells. However, it is also possible that the difference is merely due to the longer time interval between the UV survival assays on the controls and the pools/clones. The longer interval was caused by the slower growth rate of keratinocytes, in comparison to fibroblasts and CHO cells. Alternatively, the highest UV data point for the wild type keratinocytes may just be an aberrantly high value.

The B/5 PCR product in the MG#11 HAT + UV selected clones (clones 4 and 5) in figure 5.11 are interesting. The cells were transfected with *Ercc1* MG#11 and this

minigene lacks the TSS of the 1.5kb transcript. The absence of an A/5 PCR product, despite of the correct control pattern in the wild type keratinocytes, indicated that the positive B/5 reaction was not due to the presence of 1.5kb transcript. The B/5 reaction was detecting a transcript which had to be initiating upstream of the normal, 1.1kb, transcript since it contained the primer B hybridisation site. Possible explanations for this are detailed in the discussion.

Chapter 6

Ercc1 primary culture experiments

Investigation of expression of the novel Ercc1 transcript in murine tissues and primary cultures

The identification of a novel *Ercc1* transcript in the skin is intriguing due to the high biological requirement for the removal of UV-induced DNA damage in this tissue. Of all the body tissues, it is the skin that is uniquely exposed to this source of damage. It is reasonable to assume that, in order to meet the greater biological requirement for the repair of UV-induced DNA damage in the skin, *Ercc1* expression could be altered. The hypothesis is that the novel transcript confers an advantage to the skin. In order to understand the possible biological significance of the transcript in skin, it is necessary to understand the structure of the skin and the biological effects of the different types of UV. The following experiments investigate changes in *Ercc1* expression in culture, whether UV exposure could induce novel transcript expression, the effect of serum starvation upon novel transcript expression and also the difference in the size of the novel transcript produced by cells isolated from different mouse strains. In order to understand the experimental design, the following sections introduce the ideas that different wavelengths of UV are absorbed by different layers of the skin and that the cell cycle can be influenced by serum starvation to investigate expression changes due to proliferation in culture. The preliminary work leading to the various hypotheses is also explained.

The UV component of the solar spectrum

The solar spectrum and its interaction with skin is summarised by Hawk and McGregor, 2000. The ultraviolet component of the electromagnetic spectrum is the 100-400nm λ range. This range is further subdivided into three categories; UV-C is 100-290nm; UV-B is 290-320nm and UV-A is 320-400nm. DNA absorbs UV light in the region of 200-300nm (absorption peaks at 260nm), meaning that UV-C and UV-B are strongly absorbed by DNA and result in the same types of DNA damage, UV photons, being more energetic than visible light photos provide enough energy to alter the DNA chemical bonds. The UV-C component of solar radiation is completely filtered out by the ozone layer and, therefore, skin is not naturally exposed to this wavelength. Artificial sources of UV include, for example,

fluorescent lamps. Of the UV radiation that does penetrate to ground level, approximately 5% is UV-B and the remaining 95% is UV-A (with variation depending upon the time of day, year, latitude and other factors). Although the UV component of solar radiation that reached ground level is small (~5%) it is biologically important because of its absorption by DNA.

The layers of the skin

Because the skin is composed of several layers, UV radiation of different wavelengths can have different biological effects. The layers of the skin from the surface inwards are the stratum corneum, the epidermis, the dermis and the hypodermis. The stratum corneum is a layer of dead cells that move up from the underlying epidermis and is continually replaced; it protects the underlying layers of the skin from injury. The epidermis is composed of several cell types, mainly keratinocytes, and these cells gradually move up to form the stratum corneum. At the base of the epidermis are evenly scattered melanocytes which produce melanin, a UV-absorbing pigment that has a protective function. The cells of Langerhans are also present in this layer and these cells are part of the immune system (protecting against infection and tumours). Keratinocytes, melanocytes and cells of Langerhans are all nucleated cells and damage to the DNA of these cells, especially those of the basal layer, can lead to skin cancer. The underlying dermis supplies the epidermis with nutrients and oxygen via the blood vessels. The dermis contains sweat glands, hair follicles, lymphatic vessels, blood vessels and supporting fibres and the hypodermis is composed of loose connective tissue and fat.

Interaction between UV and the skin layers

Skin reflects approximately 5% of the UV that it is exposed to. The remaining 95% penetrates the tissue and is scattered then either passes out again or is absorbed by the molecules of the various skin layers. The penetrative ability of the UV is dependent upon the wavelength; short wave UV-B is mainly absorbed by the stratum corneum and epidermis, particularly by the melanin and DNA. A small fraction of the UV-B can penetrate to the dermis while longer wavelength UV-A penetrates mainly to the dermis where it is primarily absorbed by haemoglobin in the blood or

reflected back out of the skin. Absorption of UV-C or UV-B by DNA can lead to damage formation, especially pyrimidine dimers. This type of damage is very disruptive to cellular processes and hinders cell division until repaired. Therefore *in vivo*, cells of the epidermis (keratinocytes) are exposed to greatest proportion of the DNA-damaging UV-B, while the cells of the dermis (fibroblasts) are exposed to less. The most likely outcome of DNA damage by UV is complete repair by NER and most damage is repaired within hours to days. The alternative to repair is cell death. If the DNA is inaccurately repaired, permanent mutations may be introduced that could eventually lead to the development of cancer. The presence of the skin-specific transcript in skin could potentially confer an advantage in the removal of UV-induced DNA damage but if it is simply that production of the skin transcript results in higher protein levels, why could this not be achieved by simply producing more of the normal transcript?

Observations leading to hypothesis that UV could possibly induce skin transcript expression

Preliminary work in the laboratory by Dr. A. Winter had demonstrated that the ratio and expression level of normal and novel *Ercc1* transcripts in mouse skin changes dramatically in the period between day 16.5 of embryonic development and 12 weeks after birth. Until two weeks after birth the level of the normal transcript is low and constant while the level of the skin-specific transcript is very low. However, following birth the level of the skin-specific transcript increases until both the novel and normal transcripts are expressed at equal levels. The level of the skin transcript then rapidly increases while the level of the normal transcript begins to decrease until the adult expression pattern of a very high skin-specific to normal transcript is achieved (unpublished data). This change in expression of the skin-specific transcript expression is intriguing due to the possibility that the novel transcript is induced by exposure to light/UV following birth. The possibility that the skin-specific transcript could be induced by UV had to be investigated to gain further understanding of the role of the novel transcript in the skin. In order to investigate the possibility that light exposure induced skin-specific transcript production, the following experiments were designed. Primary fibroblasts isolated from various

developmental stages were protected from exposure to light and then exposed to UV/ visible light repeatedly. The resulting transcript expression pattern was then examined by means of northern blot. UV-B, UV-C and visible light were tested as possible induction agents in the following experiments. If *Ercc1* transcription was regulated by exposure to UV then transcription would be very neatly coupled to biological need for *Ercc1* activity. Induction of the skin-specific transcript by UV would particularly suggest that the reason for the skin-specific transcript's presence in the skin could be due to evolutionary selection pressure for increased UV-induced damage repair by *Ercc1* in this tissue. *Ercc1* induction experiments would more easily be performed *in vitro* than *in vivo* but the possibility of changes in *Ercc1* expression during culture required investigation.

The cell cycle response and serum starvation.

A simple description of the cell cycle in mammals is that it is divided into four discrete phases. The two fundamental parts of the cell cycle are interphase (occupies the majority of the cell cycle) and mitosis (which ends with the division of the cell). During interphase a number of growth processes occur continuously. There are also a number of stepwise processes that occur only once per cell cycle and interphase can be subdivided into three parts according to these events. The first part of interphase is referred to as G1 and this stage precedes DNA synthesis. DNA synthesis occurs during the S phase, which takes place in the middle of interphase. The stage following S phase is G2. Mitosis occurs following G2. Postembryonic cells can stop growing and leave the cell cycle for long periods (days, weeks or years); cells can enter a resting phase during G1 called G0. It is also possible for cells to enter a resting state in G2, although most non-dividing cells in the body are in G0. The cell cycle is subject to a number of growth and proliferation controls and cell division is mediated by growth factors (since uncontrolled cell division leads to tumour formation). In culture of mammalian cells these growth factors are supplied by serum and can be broad-acting growth factors or specific to particular cell types. When tissue is placed in a petri dish with serum cells migrate away from the tissue, usually fibroblasts. As serum supplies the cells with growth factors, serum starvation causes arrest in G0. Populations of somatic cells are usually not synchronous and to

study expression of genes it the cell cycle is necessary to induce synchronisation (the cell cycle and response to serum starvation is discussed in Murray and Hunt, 1993). In the following *Ercc1* cell cycle response experiment, the cells were synchronised by serum starvation and subsequent release from starvation. Reintroduction of serum permits the arrested cells to proceed with the cell cycle in relative synchrony.

Observations leading to hypothesis that expression of skin-specific transcript is dependent upon proliferation in culture

Earlier work in the laboratory by Dr. A Winter had shown that *Ercc1* transcript levels of primary dermal fibroblasts isolated from 19.5-day-old mouse embryos (E19.5) and 3-day-old mouse pups (P3) were altered by serum starvation, with the level of the skin-specific transcript decreasing in serum starved cultures and then being rapidly restored following re-introduction of serum. This observation suggested that there was possibly an *Ercc1* expression cell cycle response. Because *Ercc1* is involved in DNA damage repair and a healthy cell can respond to DNA damage by alteration of the cell cycle (e.g. arrest until damage is repaired), the possibility of a cell cycle response in *Ercc1* expression was worthy of further investigation.

Size of novel transcript in different cells

During the course of the following experiments, the observation that the skin-specific transcript expressed by cultured wild type keratinocytes (originally isolated from a BALB/c mouse), adult skin from a CD-1 mouse and dermal fibroblasts isolated from C57Bl/6 mice was of difference sizes was made. This led to sequencing of 5′ flanking region of the *Ercc1* gene in various mouse strains and the discovery that the size of the CT repeat region differed between strains.

6.1 Difference between in vitro and in vivo Ercc1 expression patterns

To investigate the possibility that culturing primary dermal and embryonic fibroblasts could lead to changes in their *Ercc1* transcript expression pattern, a northern blot was performed. To determine if the *Ercc1* expression of cultured fibroblast cells isolated from various developmental stages changed following culture, whole RNA from primary embryonic and dermal fibroblast cultures grown from tissue isolated at three different developmental stages (from C57Bl/6 mice) was compared to whole RNA from freshly dissected tissue from corresponding developmental stages (Figure 6.1 & Table 6.1). The primary cultures used for this experiment were cultured in near darkness (although they were briefly exposed to low light levels during sacrifice of the embryos/pups and separation of the dermis from the dermis).

Throughout the following analysis, the skin-specific *Ercc1* transcript (1.5kb) is referred to as the upper transcript (U) and the normal transcript (1.1kb) is referred to as the lower transcript (L). The age of embryos and pups are in days and are prefixed by the letters 'E' and 'P' respectively e.g. a 13.5 day old embryo is E13.5 while a 2 day old pup is P2. In the northern analysis tables of data are presented below the corresponding figure. The rows of the table highlighted in grey contain the data analysis of the bands on the northern following loading correction and background subtraction. This is the information that is referred to in subsequent analysis. The data presented in the non-highlighted rows of the table is the band analysis before loading correction and background subtraction and is not referred to in the analysis (provided for information only). The loading correction values are provided and the calculations are explained in the footnotes below the tables.

For the *in vitro* samples, the number of days in culture and the number of times the cells have been passaged is recorded in the figure legend (Figure 6.1). The P2 and P3 data was incomplete, due to RNA degradation, so lanes containing RNA (from an earlier northern blot) were spliced into the image. As this is a composite image, only

the upper to lower (U:L) *Ercc1* transcript ratios are examined, as this calculation is not proportional to RNA load.

If the *in vivo Ercc1* expression patterns are examined initially, to establish how the *Ercc1* expression changes during development, a trend is evident (see table 6.1); the U:L transcript ratio increases as development progresses *in vivo*. The U:L transcript ratios for the *in vivo* samples (E15 total embryo, E19 skin, P3 skin) are respectively 0.1, 0.1 and 0.4. This demonstrates a large increase in the U:L *Ercc1* transcript ratio during development. The sum *Ercc1* decreased from E15 (whole embryo) to E19 (skin) *in* vivo (this calculation is dependent upon RNA load so no figure could be provided for the P3 sample since it is from another northern). No adult C57Bl/6 mice could be spared at the time this experiment was performed so RNA from adult skin of a CD-1 mouse was included to provide an example of the final pattern of *Ercc1* expression observed *in vivo*; the high ratio of upper to lower transcript is observed in adult skin, although the U:L ratio reached in adult skin from CD-1 mice is higher than that of adult C57Bl/6 mice (Dr. A. Winter, Melton Lab., unpublished data).

If the *in vitro* samples are compared to the closest developmental *in vivo* sample then differences in *Ercc1* expression are evident. If the E13.5 *in vitro* sample is compared to the E15 total embryo *in vivo* sample, the sum *Ercc1* expression was more than 2X greater *in vitro* and the U:L transcript ratio was 3X higher *in vitro* than *in vivo.* If the E18.5 *in vitro* sample is compared to the E19 skin *in vivo* sample, the sum *Ercc1* expression was more than 4X greater *in vitro* than *in vivo* and the U:L transcript ratio was 12X greater *in vitro.* If the P2 *in vitro* sample is compared to the P3 *in vivo* sample then the U:L transcript ratio was more than 100X greater *in vitro* than *in vivo* (the sum *Ercc1* calculation is dependent upon RNA load so no figure could be provided for the P2 and P3 samples since they are from other northern blots). The *in vivo* adult skin (CD-1) showed the typical high U:L transcript ratio observed in adult CD57Bl/6 adult skin (Dr. Winter, Melton lab. unpublished data) and the ES cells had the expected low U:L transcript ratio. In summary, the *Ercc1* expression *in vitro* was different from the corresponding *in vivo* pattern as both the U:L *Ercc1* transcript ratios and sum *Ercc1* expression levels were higher *in vitro* than *in vivo* for corresponding samples.

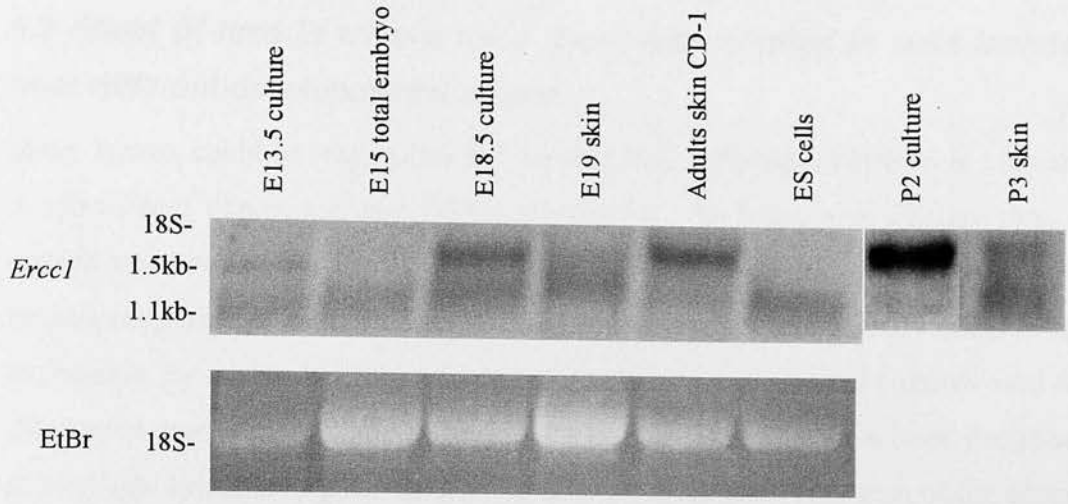
Figure 6.1 Difference between *in vitro* and *in vivo Ercc1* expression patterns in murine primary cultures and tissues

Northern blot analysis of RNA extracted from primary cultures and tissues of mouse line C57Bl/6 at different developmental stages. Upper panel, total RNA (30μg) was prepared, electrophoresed on a 1.4% agarose-formaldehyde gel, transferred onto a genescreen membrane and probed with an 800bp *BamHI* fragment of the *Ercc1* cDNA, corresponding to exons 1 to 8. The position of the 18S rRNA is indicated. Cultures isolated from macerated E13.5 embryos, E18 dermis and P2 dermis. The E13 culture had been passaged 3 times and cultured for 45 days. The E18.5 culture had been passaged 4 times and cultured for 25 days. Lower panel, ethidium bromide staining to illustrate RNA loads. The P2 culture had been cultured for 33 days and passaged 3 times. Because an adult skin sample from a C57Bl/6 mouse was not available, the adult skin RNA is from a CD-1 mouse (provided by Dr. Winter, Melton Lab.). ES cells are derived from mouse strain 129/Ola.

Table 6.1 Analysis of *Ercc1* bands in figure 6.1

Analysis of the *Ercc1* transcript bands from each lane of the northern blot. Each column of the table corresponds to the lane of the figure (northern blot and EtBr staining) directly above. Analysis was performed using Image J software.

* As P2 and P3 lanes are from separate northern blots only the U:L transcript ratio can be compared.



U before correction	838	304	4447	241	9259	1361	4334	178
L before correction	3063	5229	3580	2585	585	9154	104	449
EtBr 18s	1715	6175	4831	7660	3443	1903	*	*
Loading correction[1]	1.0	0.3	0.4	0.2	0.5	0.9	*	*
U corrected	838	84	1579	54	4612	1227	*	*
L corrected	3063	1452	1271	579	291	8251	*	*
Ratio[2] of U:L	0.3	0.1	1.2	0.1	15.8	0.1	42.7	0.4
Sum[3] U+L	3901	1537	2849	633	4903	9478	*	*

[1] Loading variation correction values calculated with first lane being assigned a value of 1.
[2] Ratio of upper to lower transcripts following background subtraction and loading correction
[3] Sum of upper and lower transcripts following background subtraction and loading correction

6.2 Effect of time in culture upon Ercc1 transcription in cells isolated from different developmental stages.

Many factors could be responsible for the observed differences between *in vivo* and *in vitro Ercc1* expression described in section 6.1. To investigate whether time in culture was responsible for the changes the following experiment was performed. Because exposure of cultured cells to ambient light could hypothetically effect *Ercc1* expression, by means of UV or visible light induction, the primary cultures used for this experiment were cultured in near darkness (although they were briefly exposed to low light levels during sacrifice of the embryos/pups and separation of the dermis from the epidermis). Following isolation, the murine primary cultures grow vigorously but in the ensuing weeks in culture the growth rate decreases until it eventually the cells stop dividing.

The cells used for this experiment were isolated from C57Bl/6 mice. Every effort was made to minimise light exposure of cells during isolation and culture so that any observed changes in *Ercc1* transcript expression could be attributed to culture conditions other than ambient light exposure. For each culture an extra dish was plated and this 'viewing dish' was cultured synchronously with the experimental dishes to monitor cell density during the experiment, minimising exposure to microscope light. See method 2.6.26 for other precautions taken to minimise light exposure.

See figure 6.2 and table 6.2. The number of passages and days in culture are shown for each sample above the figure. If the E13.5 primary culture samples are examined, the ratio of U:L transcript increased with progression of time in culture and the sum *Ercc1* expression also increased. In the E18.5 primary culture, the U:L transcript ratios increased with time in culture while the sum *Ercc1* transcription level fluctuated over time. The U:L transcript ratio of the P2 culture increased with progression of time in culture. As the 33 day P2 sample is inserted from another northern image only the ratio can be compared as the total sum *Ercc1* calculation is dependent upon RNA load.

In summary, U:L transcript ratio increased with progression of time in culture for cells isolated from all three developmental stages. There was an increase in the sum *Ercc1* transcription levels with progression of time in culture in the E13.5 sample. The sum *Ercc1* level in the E18.5 culture fluctuated but both the 15 and 16-day cultures had lower *Ercc1* expression than the 9-day culture. The changes observed over time in culture occurred in conditions of minimal light exposure.

Figure 6.2 Effect of time in culture upon *Ercc1* transcription in cells isolated from different developmental stages and cultured in near darkness.

Northern blot analysis of RNA extracted from primary cultures of mouse line C57Bl/6. Cultures from macerated E13.5 embryos, E18 dermis and P2 dermis were cultured in near darkness. RNA was isolated at various times during culture. Upper panel, Total RNA from each culture (30μg) was prepared, electrophoresed on a 1.4% agarose-formaldehyde gel, transferred onto a genescreen membrane and probed with an 800bp *BamHI* fragment of the *Ercc1* cDNA, corresponding to exons 1 to 8. The position of the 18S rRNA is indicated. Lower panel, ethidium bromide staining to illustrate RNA loads.

Table 6.2 Analysis of *Ercc1* bands in figure 6.2

Analysis of the *Ercc1* transcript bands from each lane of the northern blot. Each column of the table corresponds to the lane of the figure directly above (northern blot and EtBr staining). Analysis was performed using Image J software.

* Due to a sample problem, the P2 33-day sample in lane 7 is spliced in from another northern. The sample spliced in is the same as the one that was to be presented in the original panel (from an earlier northern blot). It is, therefore, directly comparable to the 10 day old P2 sample.

Stage	E13.5		E18.5			P2		Ker +/+ control
No. days in culture	14	27	9	15	16	10	33	
Passage no.	2	3	1	3	3	1	3	

U value Pre-correction	475	2494	8416	10349	11553	2768	12799	12820
L value Pre-correction	5683	11133	15067	12032	8472	9218	1550	0
EtBr 18S	2010	2799	3911	5195	4243	4396	3217	3187
Loading Correction[1]	1.4	1.0	0.7	0.5	0.7	0.6	*	0.9
U corrected	662	2494	6024	5576	7622	1763	*	11260
L corrected	7916	11133	10785	6483	5589	5870	*	0
Ratio of U:L[2]	0.1	0.2	0.6	0.9	1.4	0.3	8.3	All U
Sum[3] U + L	8578	13627	16810	12059	13211	7633	*	11260

[1] Loading variation correction values calculated with second lane being assigned a value of 1.
[2] Ratio of upper to lower transcripts following background subtraction
[3] Sum of *Ercc1* upper and lower transcripts corrected for loading variation

6.3 Comparison of E18.5 cultures maintained in darkness to cultures exposed to visible light or multiple UV-B or UV-C irradiation.

The primary cultures used in experiments 6.1 and 6.2 were cultured in near darkness. In order to test whether or not exposure to light could induce changes in *Ercc1* expression levels and transcript ratios, primary cultures of E18.5 murine dermal fibroblasts were cultured in near darkness and repeatedly exposed to UV-B, UV-C or visible light for short intervals. The UV-C dose delivered during each irradiation was 6 Jm^{-2}. The UV-B irradiation was performed with the UV lamp at the same height as for the UV-C irradiations and cells were irradiated for 5 minutes. The 'light control' cells were not protected from light exposure while every effort was made to minimise exposure of the 'dark control' cells to light during culture and UV irradiation was not performed on these cells. All the data sets are directly comparable as the experiment was performed using a primary culture isolated from a pool of E18.5 dermis extracted from 7 skins. The passaging, irradiations and RNA extractions were performed in synchrony. RNA samples were extracted 48h after the last UV/light irradiation.

See Figure and table 6.3. At day 9 of culture (following 1 irradiation) the sum *Ercc1* expression level and the U:L transcript ratio was higher in the dark control than in the UV-B and UV-C irradiated cultures. On day 13 of culture (following 2 irradiations) the sum *Ercc1* expression and U:L transcript ratio was higher in the UV-C irradiated sample than in the dark control and UV-B irradiated sample. The U:L transcript ratios had increased in the UV-irradiated samples but was unchanged in the dark control. Caution is required when interpreting results at later stages of the experiment because the repeated UV-C irradiations were resulting in high levels of cell mortality and the UV-B sample was also undergoing apoptosis, although to a lesser extent than the UV-C irradiated sample. This is why there is no 24 day UV-C sample and the UV-B culture was one passage behind the dark and light cultures at 24 days in culture (following 4 irradiations). By day 16 of culture (following 3 irradiations) the sum *Ercc1* expression of the UV irradiated samples and the light control was lower than that of the dark control. The U:L transcript ratio was higher in the UV-B and UV-C irradiated samples than in the dark control while the U:L

ratio in the light control was lower than in the dark control. The very high U:L transcript ratio in the UV-C culture is worthy of note (it even exceeded the ratio of the keratinocyte control by more than 3 fold). At 24 days in culture the U:L transcript ratio and the sum *Ercc1* transcription level was highest in the light control and lowest in the UV-B irradiated sample. There was no data for the UV-C irradiated sample due to total cell death.

In summary, the sum *Ercc1* expression decreased during the course of the experiment in both the dark and light controls. In the UV- irradiated samples the sum *Ercc1* expression levels initially increased following irradiation and then decreased. The U:L transcript ratio increased in all cultures. However, in the UV-B irradiated culture the increase was followed by a decrease while in the UV-C irradiated culture there was a very large increase in the U:L ratio during the course of the experiment (before the culture apoptosed completely).

Reproducibility of data in experiment 6.3: The experiment presented in figure 6.3 was performed once due to time constraints as the end of the research period approached and the associated difficulties of culturing keratinocytes for an extended period of time (increased risk of contamination). A number of northern blots were performed on the samples presented and, although these early blots were of low quality, where samples could be compared between blots the ratio of U:L transcript was consistent. Although there are dramatic ratio changes during the course of the experiment, confirmation of reproducibility would strengthen the observations. In the absence of a repeat experiment, caution is required when interpreting the results.

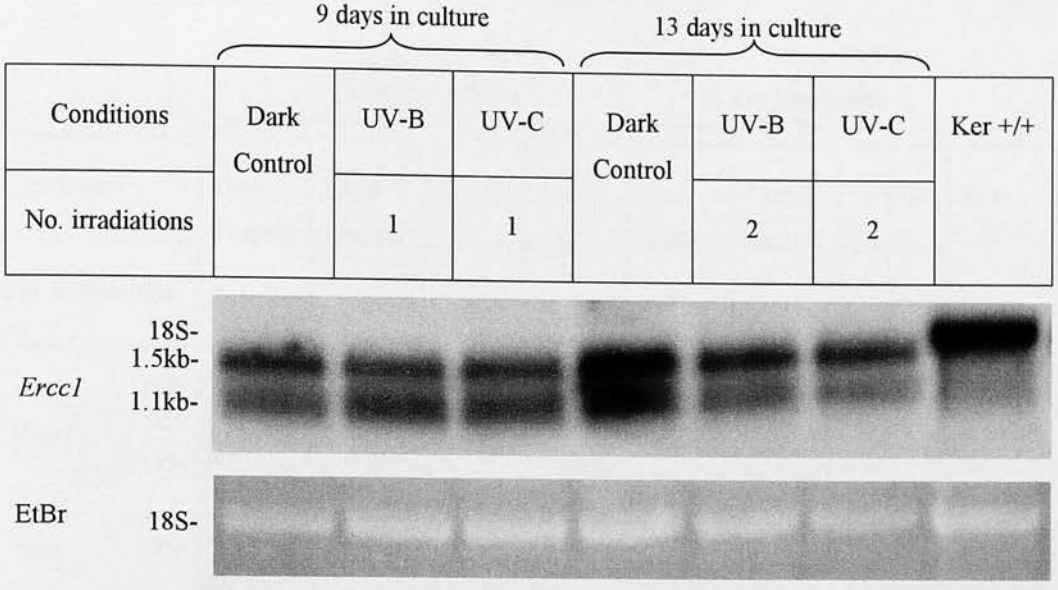
Figure 6.3 Comparison of E18.5 cultures maintained in darkness to cultures exposed to visible light or multiple UV-B or UV-C irradiations

Please note that figure and table continues over two pages.

Northern blot analysis of RNA extracted from a primary culture of C57Bl/6 E18.5 dermis which had been cultured in darkness, darkness + repeated UV-B or UV-C irradiations or light. RNA samples were taken from controls and UV irradiated samples 48h post irradiation. Upper panel, total RNA from each sample (15μg) was prepared, electrophoresed on a 1.4% agarose-formaldehyde gel, transferred onto a genescreen membrane and probed with an 800bp *BamHI* fragment of the *Ercc1* cDNA, corresponding to exons 1 to 8. The position of the 18S rRNA is indicated. Lower panel, ethidium bromide staining to illustrate RNA loads. 9-day old cultures = passage 1, 13-day-old cultures = passage 2, 16-day-old cultures = passage 3, 24-day-old cultures = passage 4 (light and dark controls, passage 3 (UV-B).

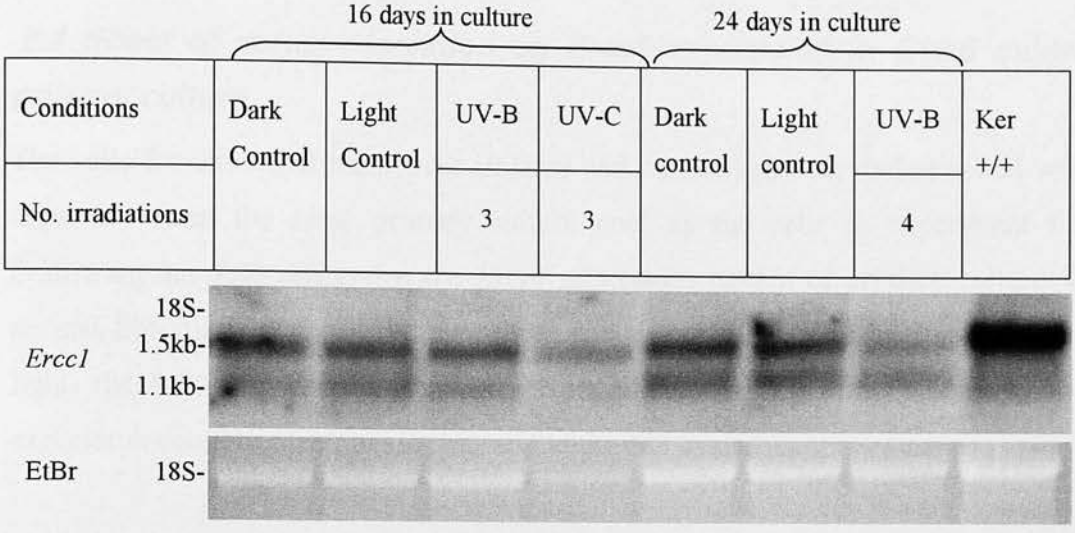
Table 6.3 Analysis of *Ercc1* bands in figure 6.3

Analysis of the *Ercc1* transcript bands from each lane of the northern blot. Each column of the table corresponds to the lane of the figure (northern blot and EtBr staining) directly above. Analysis was performed using Image J software.



EtBr 18s	2696.0	4007.0	4458.0	4395.0	3008.0	1951.0	3338.0
Loading correction values [1]	1.0	0.7	0.6	0.6	1.4	1.4	0.8
U before correction	4275.7	2411.9	2773.7	5915.4	4478.7	4545.7	16775.6
L before correction	2727.6	3837.5	3677.1	3796.0	1679.9	1052.3	611.1
U corrected	4275.7	1622.8	1677.4	3628.7	4014.1	6281.4	13549.2
L corrected	2727.6	2582.0	2223.7	2328.6	1505.6	1454.1	493.5
Ratio[2] of U:L	1.6	0.6	0.8	1.6	2.7	4.3	27.5
Sum[3] U+L	7003.2	4204.7	3901.1	5957.2	5519.7	7735.6	14042.7

[1] Loading variation correction values calculated with first lane being assigned a value of 1.
[2] Ratio of upper to lower transcripts following background subtraction and loading correction
[3] Sum of upper and lower transcripts following background subtraction and loading correction



EtBr 18s	1237	1371	2057	668	2848	2683	2294	3499
Loading correction values [1]	1.0	0.9	0.6	1.9	0.4	0.5	0.5	0.4
U before correction	4148	2263	2512	1557	3024	3473	845	16415
L before correction	1523	1146	779	16	1451	1190	1452	534
U corrected	4148	2042	1511	2883	1313	1601	456	5803
L corrected	1523	1034	469	30	630	549	783	189
Ratio[2] of U:L	2.7	2.0	3.2	97.3	2.1	2.9	0.6	30.7
Sum[3] U+L	5671	3076	1979	2913	1944	2150	1239	5803

[1] Loading variation correction values calculated with first lane being assigned a value of 1.
[2] Ratio of upper to lower transcripts following background subtraction and loading correction
[3] Sum of upper and lower transcripts following background subtraction and loading correction

6.4 Effect of serum starvation on Ercc1 expression in E18.5 murine primary culture

The cells for this experiment were isolated and cultured in near darkness and were expanded from the same primary culture pool as the cells in experiment 6.3. Following the observation that the *Ercc1* expression pattern of primary cultures of dermal fibroblasts changed during culture and that this occurred in the absence of light, the following experiment was performed to investigate the possibility that expression could be changing in response to the cell cycle.

Previous work in the laboratory by Dr. A. Winter (unpublished data) had demonstrated that the *Ercc1* expression levels a primary dermal murine culture changed in response to low serum levels in the growth media. To further investigate this observation, a serum starvation experiment on a primary culture of E18.5 dermis was performed. Cells grown under optimal serum conditions were compared to cells that had been subjected to very low serum levels for 48 hours before being released from serum starvation by the application of fresh medium of optimal serum level. Fresh medium was also added to the control dish at this time. RNA samples from the non-starved and the serum-starved cultures were taken during the course of the experiment. The final RNA sampling took place before the control group became confluent, ensuring that any transcriptional changes could be attributed to the effect of serum starvation rather than confluence level.

Please see figures 6.4, 6.5 and table 6.4. The U:L *Ercc1* transcript ratio of the control is almost double that of the serum starved sample at the 0h time point (immediately following release from serum starvation). As time progresses there is an initial decrease in the U:L transcript ratio of the starved culture but this is followed by a very rapid increase in the U:L transcript ratio (by 3 hours post release), reaching a maximum 6 hours post release. See additional figure 6.6. Between 0h and 24h post release from serum starvation the sum *Ercc1* level of the serum starved culture almost doubles, reaching a maximum at 24h. Between 24h and 48h post release from serum starvation there is a decrease in the U:L transcript ratio and sum *Ercc1* in the serum starved culture. However, the final values for the U:L transcript

ratio and sum *Ercc1* 48h post release are higher than those of the control at this time. The U:L transcript ratio of the control at 48h is half the value at 0h, so the ratio decreases in the control during the course of the experiment while the sum *Ercc1* level increases.

In summary, following release from serum starvation there was a decrease in the lower transcript production and a large increase in upper transcript production with the highest U:L transcript ratio 6h post release and highest *Ercc1* production 24h post release. The levels of the U:L transcript and sum *Ercc1* transcript remain higher in the serum starved sample than in the control sample during recovery from serum starvation. Therefore, the post-serum starvation expression level of the skin transcript is higher than the pre-serum starvation expression level.

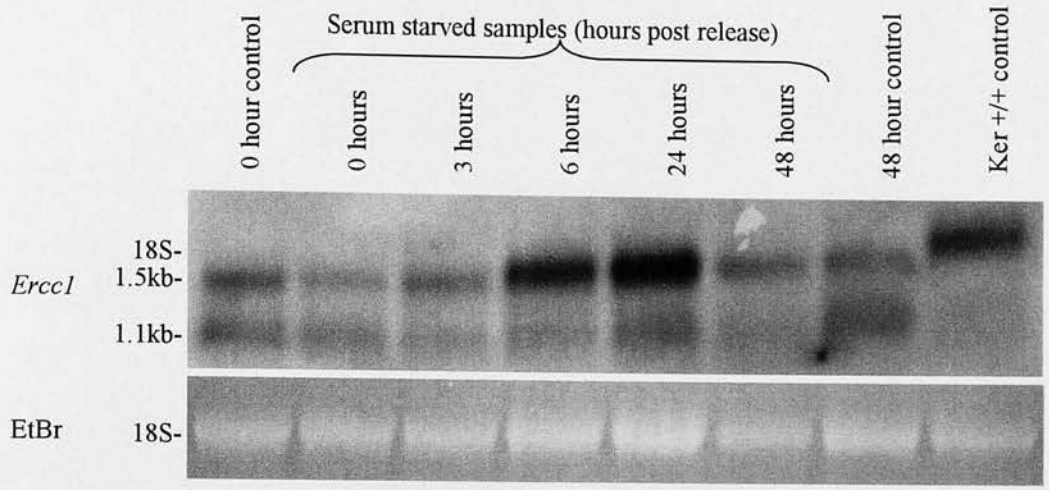
Reproducibility of data in experiment 6.4: The changes in *Ercc1* expression levels reported in experiment 6.4 are consistent with the work of Dr. Winter in the Melton lab. Serum starvation experiments not presented in the thesis were performed on E18.5 cells isolated earlier in the research time and, although the quality of the northern blots was lower, the changes in expression following release from serum starvation occur in the same order. The correlation in expression changes in the earlier experiments performed during the research time and those of Dr. Winter suggest that the changes in the level of expression and ratio of U:L transcript are reproducible.

Figure 6.4 Effect of serum starvation on *Ercc1* expression in E18.5 murine primary culture maintained in darkness

E18.5 cells were cultured in low serum for 48 hours. After 48 hours fresh medium of the optimal serum level was applied to all cultures (fresh medium was also applied to control dishes at this time). RNA samples were extracted at intervals post release from serum starvation. The control sample is from the same culture as the starved sample. Two samples are shown at 0 hours; the control has not been starved whereas the starved sample has been serum starved for 48 hours. Upper panel, total RNA from each sample (15μg) was prepared, electrophoresed on a 1.4% agarose-formaldehyde gel, transferred onto a genescreen membrane and probed with an 800bp *BamHI* fragment of the *Ercc1* cDNA, corresponding to exons 1 to 8. The position of the 18S rRNA is indicated. Lower panel, ethidium bromide staining to illustrate RNA loads. The culture was 12-days-old and had been passaged 2 times when released from serum starvation.

Table 6.4 Analysis of *Ercc1* bands in figure 6.4

Analysis of the *Ercc1* transcript bands from each lane of the northern blot. Each column of the table corresponds to the lane of the figure (northern blot and EtBr staining) directly above. Analysis was performed using Image J software.



EtBr 18s	3726	3913	3871	6039	7866	3402	4703	2653
Loading correction values [1]	1.0	1.0	1.0	0.6	0.5	1.1	0.8	1.4
U before correction	4639	2692	5015	14249	19796	5706	2863	10482
L before correction	4599	4761	1852	1890	3693	3253	5949	1680
U corrected	4639	2563	4828	8791	9377	6250	2268	14720
L corrected	4599	4534	1783	1166	1749	3562	4713	2359
Ratio[2] of U:L	1.0	0.6	2.7	7.5	5.4	1.8	0.5	6.2
Sum[3] U+L	6238	7097	6610	9957	11126	9812	6981	17080

[1] Loading variation correction values calculated with first lane being assigned a value of 1.
[2] Ratio of upper to lower transcripts following background subtraction and loading correction
[3] Sum of upper and lower transcripts following background subtraction and loading correction

Figure 6.5 Plot of the changes in *Ercc1* U:L transcript ratios in E18.5 following release from serum starvation.

Plot of *Ercc1* U:L transcript ratios in E18.5 presented in table 6.4. E18.5 cells were cultured in low serum for 48 hours. After 48 hours fresh medium of the optimal serum level was applied to all cultures (fresh medium was also applied to control dishes at this time). RNA samples were extracted at intervals post release from serum starvation. The control sample is from the same culture as the starved sample. Two samples are shown at 0 hours; the control has not been starved whereas the starved sample has been serum starved for 48 hours. The culture was 11-days-old and had been passaged 3 times at the start of this experiment.

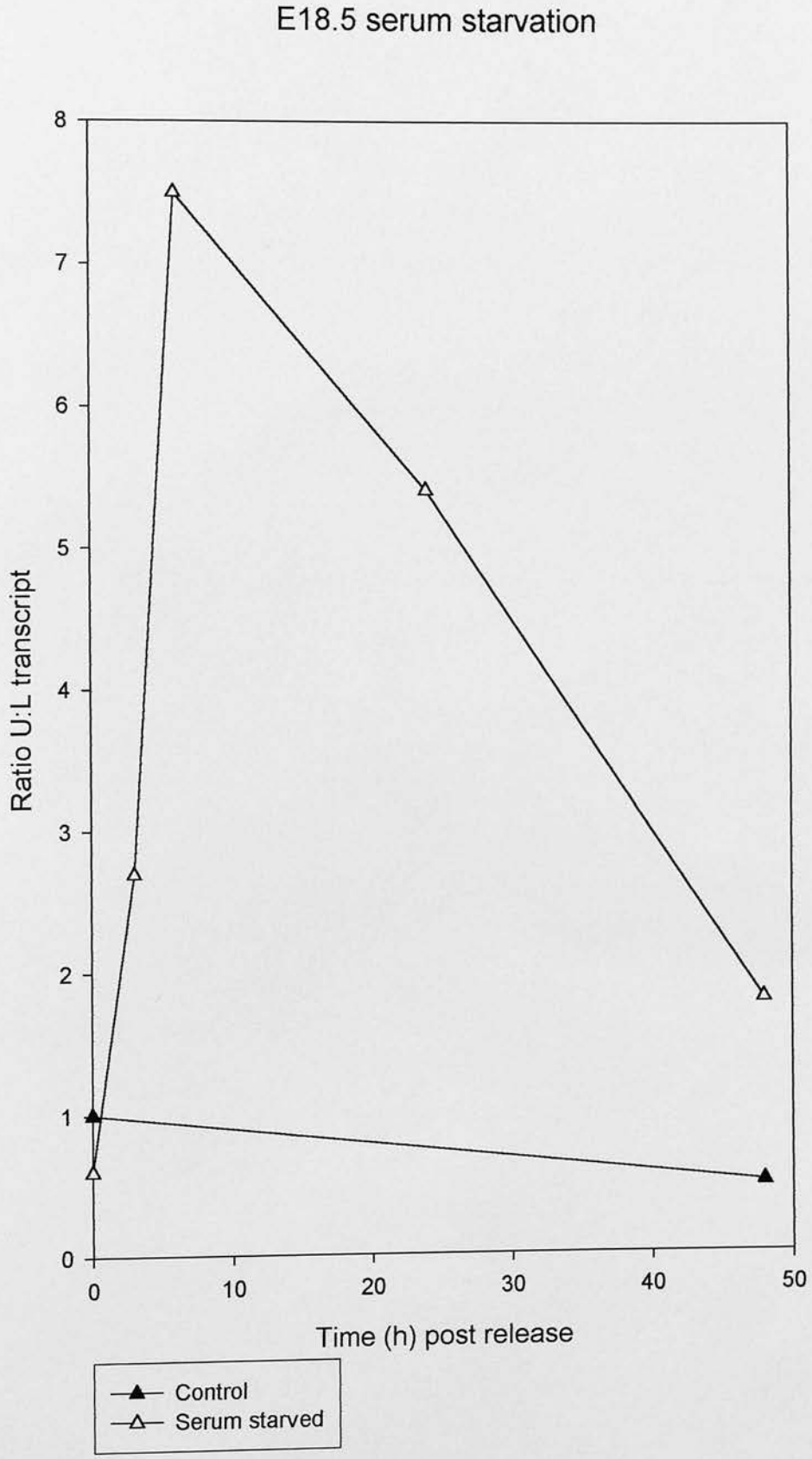
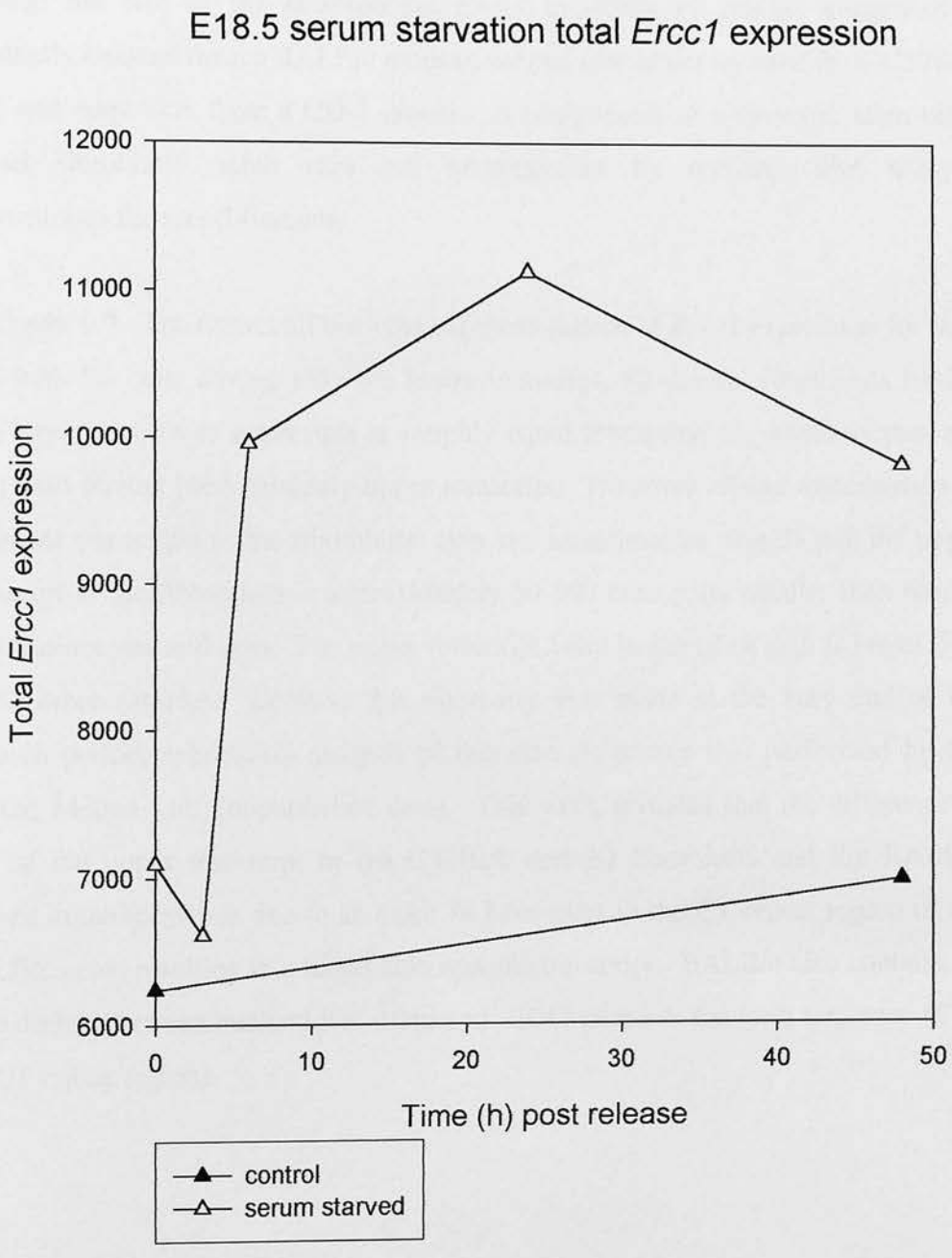

E18.5 serum starvation
Ratio U:L transcript
0
1
2
3
4
5
6
7
8
0
10
20
30
40
50
Time (h) post release
Control
Serum starved

Figure 6.6 Plot of the changes in sum *Ercc1* transcript levels in E18.5 following release from serum starvation.

Plot of sum *Ercc1* (U+L) transcript ratios in E18.5 presented in table 6.4. E18.5 cells were cultured in low serum for 48 hours. After 48 hours fresh medium of the optimal serum level was applied to all cultures (fresh medium was also applied to control dishes at this time). RNA samples were extracted at intervals post release from serum starvation. The control sample is from the same culture as the starved sample. Two samples are shown at 0 hours; the control has not been starved whereas the starved sample has been serum starved for 48 hours. The culture was 11-day- old and had been passaged 3 times at the start of this experiment.


E18.5 serum starvation total *Ercc1* expression
12000
11000
10000
9000
8000
7000
6000
Total *Ercc1* expression
0
10
20
30
40
50
Time (h) post release
control
serum starved

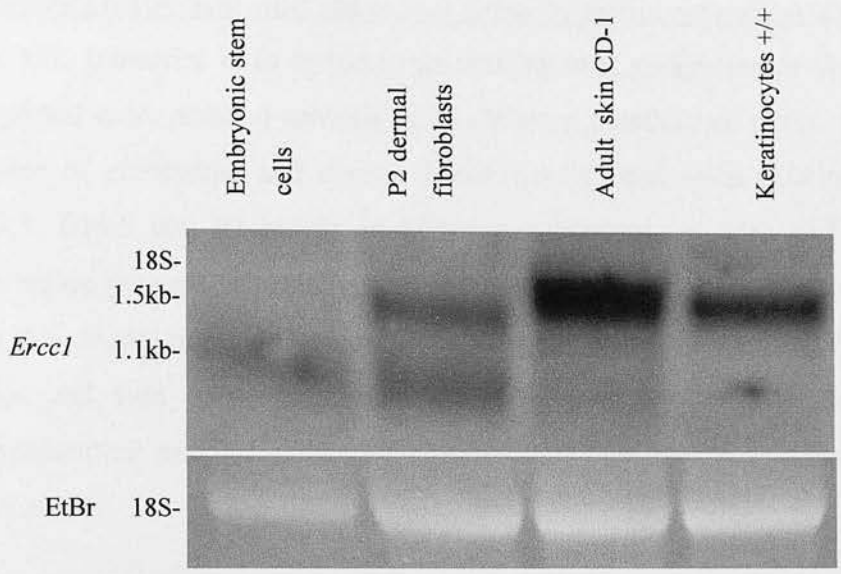
6.5 Difference in Ercc1 upper transcript size in murine dermal fibroblasts and epidermal keratinocytes

During the course of the experiments it became clear that there was a difference between the size of the skin-specific *Ercc1* transcript of murine keratinocytes (originally isolated from a BALB/c mouse), dermal fibroblasts isolated from C57Bl/6 mice and adult skin from a CD-1 mouse. A comparison of embryonic stem cells, dermal fibroblasts, adult skin and keratinocytes by northern blot analysis demonstrates the size difference.

See figure 6.7. The tissues all have the expected pattern of *Ercc1* expression for their type with ES cells having only the lower transcript, P2 dermal fibroblasts having both upper and lower transcripts at roughly equal levels and the keratinocytes and adult skin having predominantly upper transcript. However, closer examination of the upper transcripts in the fibroblasts, skin and keratinocytes reveals that the upper transcript in the fibroblasts is approximately 50-100 base pairs smaller than that of the keratinocytes and skin. The upper transcript band in the adult skin is larger than in the other samples. Because this discovery was made at the very end of the research period, subsequent analysis of this size difference was performed by Dr. Winter, Melton Lab. (unpublished data). This work revealed that the difference in size of the upper transcript in the C57Bl/6 derived fibroblasts and the BALB/c derived keratinocytes is due to an extra 74 base pairs in the CT repeat region of the BALB/c gene, resulting in a larger skin specific transcript. BALB/c also contains an extra 183bp between nucleotides –1046 and –1047 of the 5′ flanking sequence (5′ to the CT repeat region).

Figure 6.7 Difference size of upper *Ercc1* transcript in murine dermal fibroblasts and epidermal keratinocytes.

Northern blot of murine ES cells, P2 dermal fibroblasts, adult skin and wild type keratinocytes. Upper panel, total RNA was prepared from each sample (30μg), electrophoresed on a 1.4% agarose-formaldehyde gel, transferred onto a genescreen membrane and probed with an 800bp *BamHI* fragment of the *Ercc1* cDNA, corresponding to exons 1 to 8. The position of the 18S rRNA is indicated. Lower panel, ethidium bromide staining to illustrate RNA loads.



* N.B. during interpretation of upper transcript size difference that gel is smiling slightly

6.6 Conclusions from Ercc1 primary culture experiments

Difference between *in vitro* and *in vivo Ercc1* transcript expression (experiment 6.1)

The U:L transcript ratio increases as development progresses *in vivo* (observation supported more detailed analysis by Dr. Winter, unpublished data). Proliferation in culture of embryonic and dermal fibroblasts isolated from developmental stages E13.5, E18.5 and P2 results in differences between *in vitro* and *in vivo Ercc1* expression patterns, especially with regard to the skin-specific transcript production and this occurs in conditions of minimal light exposure. The U:L *Ercc1* transcript ratios and sum *Ercc1* expression levels were higher *in vitro* than *in vivo* for corresponding samples, indicating culture-related changes in skin-specific transcript expression.

Effect of time in culture upon *Ercc1* transcription in cells isolated from different developmental stages (experiment 6.2).

The U:L transcript ratio increases with time in culture in dermal fibroblasts isolated from E13.5, E18.5 and P2. There was an increase in the sum *Ercc1* transcription levels with progression of time in culture in the E13.5 and P2 cultures but the E18.5 culture did not share this trend. These changes occurred in conditions of minimal light exposure and indicated culture-related changes in the skin-specific transcript expression. The results also indicated that cells isolated from developmental stages at which the skin-specific transcript is not normally expressed at high levels have the ability to express the skin-specific transcript at higher levels.

Effect of light or multiple UV-B or UV-C irradiations upon *Ercc1* expression of E18.5 in comparison to culture maintained in darkness (section 6.3)

A dramatic increase in the U:L transcript ratio was observed in the UV-C irradiated culture during the course of the experiment. This could indicate UV-C induction of the skin-specific transcript. The peak U:L transcript ratio of the light control and UV-B irradiated samples (2.9 and 3.2 respectively) were only marginally greater than the peak U:L ratio achieved by the dark control (2.7). This suggests that neither ambient light nor UV-B induce novel transcript production. The U:L transcript ratio initially increased in the dark and light controls and the UV-irradiated samples then

decreased (except in the UV-C culture for which no data is available for 24 days). The sum *Ercc1* expression levels of the UV-B and UV-C irradiated samples initially increased following the first irradiations and then decreased while the sum *Ercc1* expression decreased during the course of the experiment in both the dark and light controls. The results may be indicative of a wavelength-dependent response to UV irradiation.

Effect of serum starvation upon *Ercc1* expression

Following release from serum starvation there was a decrease in the L transcript production and a large increase in upper transcript production with the highest U:L transcript ratio 6h post release and highest *Ercc1* production 24h post release in the serum starved sample. The U:L transcript ratio and the sum *Ercc1* transcript remained higher in the serum-starved sample than in the control sample during recovery from serum starvation. The U:L transcript ratio of the non-starved control decreased during the course of the experiment while the sum *Ercc1* level increased. These results indicated possible cell-cycle regulation of *Ercc1* skin-specific transcript production.

Size difference between novel skin-specific transcript expressed by cells isolated from different mouse strains.

The size difference between the skin-specific transcript expressed by fibroblasts (derived from C57Bl/6) and keratinocytes (derived from BALB/c) was demonstrated by means of northern blot. Later work by Dr. Winter (unpublished data) found that this size difference is due to a strain difference between the mice from which the cells were isolated. An extra 74 bps are present in the CT repeat region of the *Ercc1* gene's 5′ flanking region in BALB/c.

Chapter 7

Discussion

The aims of the work presented in this thesis were firstly to investigate the role of sequences present in the 5′ flanking sequence of the *Ercc1* gene in expression of the skin-specific transcript and secondly to investigate the role of the skin-specific transcript in the skin. The novel, skin-specific transcript is the predominant transcript in the skin. The normal *Ercc1* transcript is expressed constitutively at low levels throughout the body. This led to the hypothesis that the skin-specific transcript may have biological significance since the Ercc1 protein is part of the NER pathway, the primary pathway for removal of UV-induced DNA damage, along with other categories of DNA damage. Since out of all the body tissues, the skin is uniquely exposed to UV and consequently has the highest biological demand for repair of this type of damage the expression of the skin-specific transcript at high levels and at a high ratio to the normal transcript in the tissue is intriguing. The presence of the skin-specific transcript could indicate that this transcript confers an advantage to the skin in helping to meet the biological demand for repair in this tissue. The coding region of both the skin-specific and normal transcripts is the same size and the expression level of the transcript in the skin is directly proportional to protein level (Dr Winter, Melton Lab., unpublished data). The relationship between transcript level and protein production is interesting because if the purpose of the high expression level of the skin-specific transcript is merely to increase the level of Ercc1 protein in the skin then why could this not be achieved more simply by up-regulation of normal transcript production? Since the skin-specific transcript is expressed at high levels in the skin and is the predominant transcript in this tissue, there may be something unique about the control of expression of this transcript in the skin. The skin-specific transcript would, therefore, have a role in the skin that cannot be performed by the normal transcript, i.e. upregulation of expression in response biological demand or a tissue-specific expression signal. These possibilities are addressed during this discussion. The original hypothesis behind the work presented in this thesis was that the skin-specific transcript conferred an advantage to the skin in meeting the biological demand for repair of UV-induced damage in this tissue.

Gene expression can be constitutive and when this is the case the gene is expressed all the time in all tissues. This is how the normal transcript is expressed in the body, constitutively at low levels. In contrast, gene expression may be tissue-specific or induced by environmental stimuli. The control of gene expression in eukaryotes was described in the introduction to chapter 3 (section 3.1). The work presented in chapters 3, 4 and 5 was performed to investigate the importance of regions in the 5′ flanking region of the *Ercc1* gene to expression of the skin-specific transcript. The *Ercc1* gene's 5′ flanking region contains none of the classical promoter elements such as TATA, CAAT or GC boxes (aside from a TATA box and a GC box that are very distal from the TSS of either the skin-specific or normal transcript and unlikely to be functional). This is in keeping with the situation in other NER genes as the lack of classical promoter elements in the promoter regions of such genes had been previously reported. The promoter of the normal human *ERCC1* transcript was characterised and found to reside within 170bp of the normal transcript's TSS (van Duin *et al.*, 1987). This 170bp fragment of 5′ genomic DNA was capable of driving expression of a complete human cDNA and corrected the DNA damage repair deficiency of *Ercc1* deficient CHO43-3B cells. However, the discovery of the skin-specific transcript which initiates upstream of this 170bp region suggested the likely presence of a second, skin-specific promoter. Because of the lack of classical promoter elements, other sequence elements are likely to be responsible for regulation of skin-specific transcript expression. A number of interesting sequence elements have been identified in the 5′ flanking sequence of the *Ercc1* gene, including the CT repeat and a more distal (CA) repeat (Hsia, Melton Lab. unpublished data). The CT repeat, due to its proximity to the skin-specific transcript TSS and the role of such regions in regulation of expression in other genes, was chosen for investigation as possibly being involved in skin-specific transcript expression. The use of minigenes was preferable over other methods of investigation of the skin-specific promoter for a number of reasons. Firstly, as explained in section 3.1, the likely presence of two promoters precluded the use of reporter genes for characterisation of the skin-specific promoter. This was because the presence of the normal promoter would interfere with detection of reporter gene product resulting from the skin-specific promoter activity.

Both the murine *Hprt* and *Ercc1* loci, having respective spans of 33kb and 15kb, are too large for standard cloning and transfection to be successful. In contrast, the *Hprt* minigene, *PGKHprt*, spans a region of only 2.7kb and contained convenient restriction sites that enabled linearisation prior to transfection without disruption of the coding region. This made this minigene an ideal selectable marker for the transfections. In addition to the problem of the size of the murine *Ercc1* locus, the region was not available as a single cloned genomic fragment. Minigenes are a better choice for transfection than large cloned fragments. Three minigenes had been constructed; *Ercc1* MG#13 was 4.4kb smaller than the *Ercc1* locus while the deletions in *Ercc1* MG#11 and #20 resulted in these constructs being even more compact than MG#13. Identification of cell lines which expressed the skin-specific transcript, as well as cell lines which did not normally express the skin-specific transcript, alongside *Ercc1/Hprt* null corresponding cell lines provided an in vitro environment for performing deletion analysis using the *Ercc1* minigenes.

The *Ercc1* minigenes #11, #13 and #20 were available for investigation of regulation of skin-specific transcript expression. A transfection experiment involving *Ercc1* MG#13 had previously been performed by Dr. Jim Selfridge (Melton Lab.) and this demonstrated that the MG corrected the UV hypersensitivity of *Ercc1*-deficient fibroblasts, CHO cells and keratinocytes. This earlier experiment differed from those presented here in that the analysis of transcript production was by means of northern blot rather than RT-PCR. RT-PCR was preferable over northern blot analysis because it can be performed more quickly and does not involve the use of radioactivity Size differences between transcripts due to differential polyadenylation would not prevent identification of the two different transcripts by RT-PCR either. The *Ercc1* MG#13 work demonstrated that the 5' flanking region included in this minigene contained the elements necessary for skin-specific transcript expression. *Ercc1* MG#11 and #20 presented new opportunities for more detailed investigation of the *Ercc1* expression patterns and skin-specific transcript production.

Details of how the minigenes were constructed are presented in Chapter 3. To summarise, the three minigenes were constructed from a cloned genomic fragment,

pKpnI 9.3 (containing exons 1-7) ligated to the 3′ end of the *Ercc1* cDNA (containing exons 7-10) and a short 3′ UTR that contained the polyadenylation signal. These parts were joined to the 5′ flanking sequence of the *Ercc1* gene that was likely to contain the skin-specific transcript's promoter. However, the CT repeat region, possibly part of the skin-specific transcript's promoter, was deleted from the flanking sequence in *Ercc1* MG#20. The entire sequence upstream of a *Kpn1* restriction site in the flanking sequence, including the skin-specific transcript's TSS and the region likely to contain the promoter, was deleted in *Ercc1* MG#11. In this way, deletions in the flanking sequence enabled the importance of the missing sequences to skin-specific transcript expression to be ascertained by performing RT-PCR, while the functionality of the minigenes was demonstrated by the degree of correction of UV hypersensitivity of *Ercc1* deficient cells in the UV survival assays.

The three minigenes allowed a number of issues to be addressed:

5. Were all three minigenes functional in the three different cell lines, including that of the non-murine species?

6. Were the three minigenes expressed in a manner appropriate to the cell type transformed?

7. Assuming expression of *Ercc1* MG#13 was appropriate for the cell type, did the deletions in *Ercc1* MG#11 and #20 result in altered transcript expression patterns/UV survival when compared to cells expressing *Ercc1* MG#13?

The minigene transformations of the three different cell types were performed consecutively; the fibroblast experiments were followed by the CHO experiments and finally the keratinocyte experiments. Alterations to the experimental design in areas such as plating density were made according to the growth rate and requirements of the cell type. Where new possibilities for investigation of skin-specific transcript production arose, additional features were added to run in tandem with the main experiments. This allowed such issues as the possible impact of the

selection methods post electroporation upon clonal expression patterns and whether or not the selected clones were representative of transformed cells to be addressed in the later experiments.

The three minigenes were used for electroporation of three cell types with varied *Ercc1* expression patterns in wild type counterparts. The minigenes permanently integrate into the genome of the transformed cells. The three cell lines were chosen for their *Ercc1* and *Hprt* null status. The *Ercc1* and *Hprt* deficient murine embryonic fibroblast cell line ($PF24^{+}$#2D-4) provided an environment for testing whether transfection would result in production of both transcripts, as occurs in wild type fibroblasts. Fibroblasts were used in the early experiments because they have a similar pattern of expression to skin (expressing both transcripts but with a lower skin-specific to normal transcript ratio). Fibroblasts also have faster growth rate than keratinocytes and are much less technically demanding to culture. Transfection of *Ercc1* and *Hprt*-deficient CHO cells, CHO43.3B, was performed to demonstrate the species specificity of the minigenes and whether they would be expressed appropriately in a cell line that normally only expresses the 1.1kb transcript. The final cell type used for transfection was *Ercc1* and *Hprt* deficient murine keratinocytes. Wild type keratinocytes have a high skin-specific (1.5kb) to normal (1.1kb) transcript ratio, the expression pattern observed in skin. As long as the expression pattern of *Ercc1* MG#13 was correct for the cell type, alterations in levels of the skin-specific transcript between cells expressing *Ercc1* MG#13, #11 and #20 could be attributed to the deletions in those minigenes, indicating whether deleted regions formed part of the skin-specific transcript's promoter or enhancer regions and whether lack of skin-transcript affected correction of UV-hypersensitivity.

The UV survival assays differed slightly between the three cell lines. The initial plating density was determined by the growth rate of the cells. CHO cells, having the greatest rate of growth, were plated at a lower density than fibroblasts while keratinocytes were plated at a higher density. The UV range used was determined by the survival curves of the wild type and UV hypersensitive lines. The highest UV dose needed to be such that the wild type cells showed little, if any, cell death and the

doses were clustered where the most useful information could be obtained. Since wild type keratinocytes originate from the skin they are very UV resistant when compared to embryonic fibroblasts and CHO cells that would not in their native state be exposed to UV. This meant the UV doses received by the keratinocytes were higher than in the fibroblast and CHO UV assays. The keratinocytes required different trypsinisation conditions as they adhere strongly to the dish surface. Plating conditions were also different because these cells do not grow well at low density (meaning that expansion of clones post transfection was complicated). Additional experiments on selection method post electroporation and the representative nature of the clones were performed on keratinocytes and these results are discussed after the UV and RT-PCR clone work in the three cell lines.

Following electroporation and selection, the ratio of HAT+UV selected colonies to HAT only selected colonies for the minigenes in each cell type (table 3.2, 4.1 and 5.1) demonstrated that the presence of the *Ercc1* minigene conferred resistance to UV because the ratio was much higher in the minigene electroporations than in the *pBT/PGK-HPRT* only electroporations. In the fibroblast and keratinocyte transformations (table 3.2 and 5.1), the ratio HAT+UV selected colonies to HAT selected colonies was lower in the *Ercc1*MG#11 transfection than in the *Ercc1* MG#13 and 20 transfection. This is interesting because transfection with *Ercc1*MG#11 meant the cells could not express the skin-specific transcript (which is the wild type expression pattern). Perhaps the inability to express this transcript resulted in fewer cells surviving the UV selection than in the case of the transfections with *Ercc1* MG#13 and #20 (which enabled expression of the skin-specific transcript). The ratio of HAT+UV selected colonies to HAT selected colonies was higher in the CHO cells transfected with *Ercc1* MG#13 and MG#20 than with *Ercc1*MG#11 and this could indicate that the cells transfected with *Ercc1*MG#11 were less resistant to UV-irradiation post electroporation than those expressing *Ercc1* MG#13 and #20 (table 4.1). However, without repeating the experiments with a greater number of dishes plated with cells for each transfection, no conclusions regarding the statistical significance of the ratio differences can be drawn. In results discussed later, transformation of CHO cells with minigenes #13 and #20 resulted in

production of an abnormal transcript which initiated upstream of the normal transcript but was not the skin-specific transcript. The hypothesis that the UV selection post electroporation could be selecting non-representative clones in this cell type led to alteration of experimental design in the keratinocyte experiments.

The three minigenes corrected the UV hypersensitivity of the *Ercc1*-deficient fibroblasts, CHO cells and keratinocytes in the clonal analysis and there was little difference in the level of correction between the minigenes. The only notable difference was a slightly greater variability in the UV survival of fibroblast and keratinocyte clones transfected with *Ercc1* MG#20 in comparison to *Ercc1* MG#13 (figures 3.6, 3.8, 5.1 and 5.7). The correction of UV hypersensitivity presented in chapters 3, 4 and 5 indicated that the three minigenes were being transcribed in the transfected cell lines and that functional Ercc1 protein was being produced. These results also indicated that the three minigenes were functional in both murine (fibroblasts and keratinocytes) and non-murine (CHO) cells. Although UV survival assays demonstrated that the transfected cells were producing functional protein it was also necessary to investigate which transcripts were being expressed. These results indicated whether the minigenes were being expressed appropriately for the cell type and whether the deletions in *Ercc1* MG#11 and #20 resulted in altered expression patterns in transfected cells.

The RT-PCR results for the transfections with *Ercc1* MG#13 (showed that the minigene was being expressed appropriately in the transformed fibroblasts and keratinocytes (figures 3.10, 5.10 and 5.13). This meant that all the elements required for correct expression of the skin-specific transcript were contained in the minigene. The expression pattern observed in transformed CHO cells was not appropriate for the cell type as skin-specific transcript was expressed in the clones, although the bands detecting skin-specific transcript were very faint in 3 out of 4 of the clones (so there was a semblance of the correct expression pattern as long as this is proportional to transcription level) (figure 4.5). This indicated that low levels of skin-specific transcript were being produced. Speculation upon the possibility that post electroporation UV-selection could be selecting clones that had abnormal expression

led to the use of pools of cells and a comparison of cells selected with and without UV in later experiments. Any advantage conferred by the expression of the skin-specific during the UV selection post-electroporation was not demonstrated by the UV survival assays as the presence of skin-specific transcript in the CHO cells did not improve the UV survival ability of the cells when compared to the wild type (figure 4.2). The abnormal expression could also be due to the sequence in the minigene being murine in origin or lacking in a sequence element required for correct expression in CHO cells.

The effect of the deletion in *Ercc1* MG#11 upon transcript expression needs to be considered when the RT-PCR results are analysed. Because the deletion removes the TSS of the skin-specific transcript it is not possible to express the skin transcript following transfection with this minigene, even in cell lines where the skin-specific transcript is produced in wild type cells. Transfection with the minigene resulted in production of the normal (1.1kb) transcript in fibroblasts, CHO cells and 3 of the 5 keratinocyte clones (the exceptions are discussed later) (figures 3.11, 4.6, 5.11 and 5.13). This expression patterns were as expected in the three cell types with fibroblasts and keratinocytes unable to express the skin-specific transcript due to the deletion in MG#11. The lack of skin-specific transcript in the cell types that express it in their wild type state, fibroblasts and keratinocytes, did not result in less correction of UV hypersensitivity when compared to cells transfected with *Ercc1* MG#13 expressing both transcripts.

The deletion in *Ercc1* MG#20 was made to investigate the importance of the CT repeat region for production of the skin-specific transcript. The results have to be interpreted in parallel with those of *Ercc1* MG#13 since the only difference between the two minigenes is the deletion in *Ercc1* MG#20. The RT-PCR work on *Ercc1* MG#13 demonstrated correct expression of both transcripts in fibroblasts and keratinocytes (figures 3.10 and 5.10 and 5.13). Transfection with *Ercc1* MG#20 resulted in an altered expression pattern in three out of four fibroblast clones and the normal fibroblast expression pattern in one clone (figure 3.12). Of the three clones expressing the skin-specific transcript, the level of skin-specific transcript was

greatly reduced in 2 of the clones when compared to the RT-PCR pattern observed in fibroblasts transformed with *Ercc1* MG#13 (figure 3.10). Because the RT-PCR technique used in these experiments was not quantitative it is only possible to speculate upon this observation. If the reduced brightness in the bands of the RT-PCR performed upon the cells transfected with MG#20 is proportional to the original level of transcript expressed in the cells then the CT repeat region deleted in *Ercc1* MG#20 may be important for correct expression of the skin-specific transcript in fibroblasts. However, the clones were all expressing the upper so, in a purely qualitative interpretation, there is no difference between the cells transfected with MG#13 or #20. As stated earlier, the qualitative reporting on band levels is for interest only and any observations upon differences between minigene transfection results drawn from the level of RT-PCR product bands rather than the absolute absence or presence of bands would have to be confirmed or refuted by use of quantitative RT-PCR. The deletion in MG#20 did not prevent skin-specific transcript production so this region is not absolutely crucial for skin-specific transcript production but it may be required for regulation of expression and this could be further investigated using qualitative RT-PCR. The expression pattern of *Ercc1* MG#20 in transformed keratinocytes was correct for the cell type. A difference in the level of brightness in the RT-PCR on one HAT only selected clone was observed, the level of skin-specific transcript appeared to be lower than in the other clones (5.12), but this could be an artefact of the RT-PCR reaction and would need to be further investigated by use of quantitative RT-PCR to discover if this was a true reflection of a difference in the original expression level of the transcript. As in the case of the *Ercc1* MG#11 transfections, lack of skin-specific transcript in cells that would normally express this transcript did not lead to a reduction in correction of UV hypersensitivity in comparison to cells transfected with *Ercc1* MG#13 (figures 5.1 and 5.7). The expression pattern in *Ercc1* MG#20 transfected CHO43.3B cells was incorrect for the cell type in 3 of the 4 clones (figure 4.7 as the presence of product bands detecting expression of the upper transcript demonstrated an incorrect expression pattern in the 3 clones. Although the level of band intensity for the skin-specific transcript product varied between the clones and was generally quite low, this is a qualitative interpretation and would have to be confirmed/refuted using

quantitative RT-PRC. If low levels of skin-specific transcript really are being expressed by these clones then there is a semblance of the normal pattern in comparison to the transformed fibroblasts and keratinocytes. However, it may be that quantitative RT-PCR demonstrates no correlation between the expression level in transcformed clones and the level of product reported here. Because CHO cells transfected with *Ercc1* MG#13 also expressed skin-specific transcript the incorrect expression pattern in cells transfected with *Ercc1* MG#20 is probably not be due to the CT repeat region being deleted in this cell type. The possibility that the post-electroporation UV selection regime and clonal selection was resulting in selection of non-representative clones was address in later experiments.

The observations of abnormal expression in the transformed CHO cells led to the inclusion of some additional components in the subsequent keratinocyte experiments. Transformation of CHO43.3B with *Ercc1* MG#13 and #20 had led to inappropriate expression of skin-specific transcript. This did not appear to confer a survival advantage at the stage of the UV survival assays but it was hypothesised that the UV selection regime post electroporation could be leading to cells with abnormal expression of skin-specific transcript being favoured over those with correct expression patterns; following electroporation of keratinocytes, selection by HAT and UV or HAT alone was performed to investigate this possibility. Due to the 3:1 ratio of *Ercc1* to *Hprt* DNA in the electroporation, cells surviving HAT selection were likely to also contain the *Ercc1* DNA and if the clones did not contain *Ercc1* minigene DNA then they would not exhibit corrected UV hypersensitivity. The comparison between the HAT selected and HAT + UV selected clones would reveal any difference in expression between the two sets. Because keratinocytes normally express both transcripts, selection for cells expressing skin-specific transcript at high levels would have been more striking in CHO, which normally only express normal transcript. Pools of cells were also examined and compared to the clones to see if the clones were representative of transfected populations of cells. Because a clear difference was observed between the transformed keratinocyte UV Vs non-UV selection sets the work was not repeated in CHO. Had no difference been observed it would have been necessary to repeat the work in CHO cells.

Keratinocytes clones transformed with *Ercc1* MG#13 or #20 did not demonstrate greater skin-specific transcript production in clones selected by UV post-electroporation when compared to clones selected by HAT only (figures 5.10 and 5.12). In the clones transfected with *Ercc1* MG#11, a transcript initiating upstream of the normal transcript (but not the skin-specific transcript) was present in the clones selected by UV but not HAT alone (figure 5.11). Keratinocyte pools transfected with the *Ercc1* MG#11 also had this unusual expression pattern (figure 5.13). These cells were producing the expected 1.1kb transcript but the cells that had undergone UV and HAT selection were also expressing a transcript that was detected by the B/5 reaction but not the A/5 reaction, meaning it was not the skin-specific 1.5kb transcript. There are two possibilities that could explain transcripts initiating upstream of the 1.1kb transcript in cells transformed with MG#11; transcripts were initiating somewhere between primer A and primer B's hybridisation sites or in the adjacent plasmid sequence. Because the primary transcript in keratinocytes is the skin-specific transcript there may be selective pressure for initiation of transcripts upstream of the normal transcript if this confers a survival advantage when the cells are irradiated with UV. Perhaps in the absence of the skin-specific transcript, the B/5 reaction was able to detect transcripts initiating downstream of the primer A hybridisation site. Because the B/5 products were only observed in the UV selected clones this may indicate that this is the case and that there was selection for cells expressing upstream transcripts in the post electroporation UV selection regime. If the transcript being detected by the B/5 reaction was initiating between the *Kpn1* site and the primer B hybridisation site then it would not contain the primer A hybridisation site. The skin-specific transcript is heterogeneous and not all the skin transcripts start at the 'skin' TSS. This heterogeneity could spread as far as the region between the *Kpn1* site and the primer B hybridisation site. Such transcripts can only be detected by RT-PCR if the 1.5kb transcript is missing and would reflect a pressure to initiate upstream of the normal transcript in keratinocytes. Alternatively, there could be a structural explanation for the positive B/5 reaction. The transcripts detected could be initiating in the plasmid DNA upstream of the primer B hybridisation site. In this case the transcripts would be mere artefacts. The

cells which had not been subjected to UV irradiation did not have any transcript initiating upstream of the 1.1kb transcript. The presence of the unusual transcript in the keratinocytes transfected with *Ercc1*MG#11 did not confer greater correction of UV hypersensitivity when these clones were compared to those that were expressing only the 1.1kb transcript. The lack of skin-specific transcript did not reduce the UV hypersensitivity in the keratinocytes and fibroblasts transfected with *Ercc1* MG#11, which expressed only the 1.1kb transcript. The results of transfection with *Ercc1* MG#11 are intriguing because the selection with UV post electroporation in keratinocytes appeared to result in selection of clones which were expressing any transcript initiating upstream of the 1.1kb transcript when compared to keratinocytes that had HAT selection only. This suggested there could be a selective advantage to possession of such a transcript in conditions of UV exposure.

A second experiment was running in conjunction with the keratinocyte clonal transfections. The clones chosen for expansion were representative in terms of size when compared to other clones in the dish but perhaps due to the small number of clones studied, random selection resulted in expression patterns that differed slightly from the 'average' expression pattern of transfected cells. Shifts in expression patterns could possibly be more obvious in entire populations of clones treated as pools and an 'average' expression pattern of transfected cells could be observed. For each minigene two pools were available – one selected with HAT only post electroporation and one selected with HAT + UV. There was no difference between the UV survival between transfected clones and pools (figures 5.3, 5.6 and 5.9) or between pools selected by UV and pools selected without UV (figures 5.2, 5.5 and 5.8). However, when the RT-PCR results for the pools were analysed there was a difference between the two selection methods (figure 5.13). Pools transfected with *Ercc1* MG#13 and selected by either method had the normal pattern of expression. Of the keratinocyte pools transfected with *Ercc1*MG#11, the pool that underwent HAT selection expressed only the 1.1kb transcript while the pools selected with HAT + UV was expressing a transcript detected by the B/5 reaction but not the A/5 reaction, as with the keratinocyte clones transfected with *Ercc1* MG#11. The pool transfected with MG#20 and then exposed to HAT was expressing both transcripts,

as was the pool transfected with this minigene and exposed to HAT + UV. In quantitative terms, both were the same. There were differences in the intensity of product bands for the skin-specific transcript, with the pool exposed to HAT only having much a lower intensity product band than that of the pool selected with HAT + UV. This could be an indication of selection for higher expression of the skin-specific transcript following UV exposure during selection if the intensity of the band is directly proportional to the level of transcript or it could be a mere artefact of the RT-PCR process. This would be confirmed or refuted by quantitative RT-PCR but can only be speculated upon in the absence of such data..

Assuming the difference in expression of skin-specific transcript was genuine and due to the UV component of the selection post electroporation, it was hypothesised that UV irradiation of a pool selected with HAT only should induce a shift in expression pattern (in terms of band intensity) towards that of the pool selected with HAT + UV. In the absence of follow-up experiments using quantitative RT-PCR to demonstrate whether the intensity of product band was a true reflection of the original transcript levels or merely an artefact of the reaction the following observations are merely speculative. This is because in terms of quantitative differences, the bands are present in both pools and , therefore, there is no difference between the pools . There were qualitative differences in the intensity of the bands but the RT-PCR used was not quantitative so these are reported for interest only and should not be over-interpreted. The keratinocyte pool transfected with *Ercc1* MG#20 showed the greatest difference in skin-specific transcript production between the two selection regimes in terms of band intensity, with less skin-specific transcript following HAT only selection. Following RT-PCR on RNA extracted from the HAT selected pool 24 hours after a UV dose of 6 Jm^{-2} was applied, an increase in the intensity of the skin-specific transcript product band was observed when compared to the non-irradiated sample (figure 5.14). Speculatively, this could indicate that UV-C irradiation had resulted in increased expression of the skin-specific transcript. However, in the absence of follow-up experiments using quantitative RT-PCR no conclusions may be drawn from observations upon product band intensity, only upon the absolute presence or absence of product in RT-PCR reaction. The pattern of

expression observed with keratinocyte pools transfected with MG #20 and treated with UV + HAT or HAT alone was interesting, as there was a possible suggestion that UV irradiation led to an up-regulation of upper transcript production. To study this using quantitative RT PCR would confirm that the results reported here are a reflection of actual differences in expression level rather than artefacts of the RT-PCR process.

Following my departure form the Melton Lab. the cells used in these experiments were re-tested to check the reproducibility of the results I reported with regard to MG#20 transfection and HAT or HAT+UV selection; RT-PCR performed upon RNA extracted from cells frozen during my research period (which had been transfected with MG #20 then treated with HAT + UV or HAT alone) generated the same RT-PCR results both in terms of the absence or presence of bands and the intensity level of the bands as reported here, but this was not quantitative RT-PCR. It is accurate to say that the patterns of expression visualised by the RT-PCR process presented here were consistent during the course of the experiment between clones; i.e. the same pattern of expression was obtained for any particular clone/pool as long as the cycle number and initial quantity of RNA was identical. Increased cycle number or quantity of RNA used for the RT-PCR resulted in brighter bands but this amplification was consistent across the samples and the pattern of expression was still preserved. However, without the use of quantitative RT-PCR, the observations regarding the level of transcript expression when investigated with RT-PCR presented in this thesis must be interpreted with caution and are reported for interest only. The RT-PCR technique used is not quantitative and interpreting levels following non-quantitative RT-PCR is not an appropriate use of the technique. However, the actual presence or absence of transcript detected via the RT-PCR experiments presented is an appropriate use of the technique and reproducible. Strictly adhering to reporting only the absence or presence of transcript renders the difference in 'levels' reported irrelevent, making the differences between pools of keratinocytes transfected with MG#20 and treated with HAT or HAT +UV zero.

The work in the minigenes, therefore, indicated that the 1.05kb *Ercc1* gene flanking sequence contained all the elements necessary for production of the skin-specific transcript. The work also slightly indicated that the CT repeat region may be important for correct expression of the skin-specific transcript. Deletion of this region resulted in an absence of skin-specific transcript expression in comparison the cells transfected with MG#13 in one of the fibroblast clones. Observations upon differences in product band intensity in the RTPCR analysis of the transfected clones and pools are of reduction in the intensity of skin-specific transcript product bands in cells that normally express the skin-specific transcript when transfected with MG#20 in comparison to MG#13. The work also slightly indicated that expression of the skin-specific transcript may be UV-C inducible in the keratinocyte pool transfected with *Ercc1* MG#20 (lacking the CT repeat region) but this observation is also drawn from product-band intensity differences in the RT-PCR results and, for the reasons explained, in the absence of follow-up work with quantitative RT-PCR no conclusions may be drawn from such observations and they remain mere speculation.

Further work with the minigenes could include deletion of the 170bp region that harbours the normal transcript's promoter to investigate the effect of removing the normal transcript upon correction of UV hypersensitivity of *Ercc1*-deficient cells. Deletion of the CA repeat region located upstream of the skin-specific transcript's TSS would be interesting because CA repeats can function as enhancers in the same way as CT repeats, which had been shown to be involved in skin-specific transcript production. The CT repeat region is discussed in more detail later.

In addition to the minigene deletion work for identification of regions of importance to expression of the skin-specific transcript, the role of this transcript in the skin was also investigated. The normal developmental pattern of *Ercc1* expression in C57Bl/6 mice is a gradual increase in expression of the normal transcript early in development to a final constitutive, low level of expression. In the skin, the level of skin-specific transcript rapidly increases after birth and the ratio of skin-specific to normal transcript increases until the adult pattern of a high skin-specific to normal transcript ratio is achieved. At the time the experiments presented in this thesis were planned,

the hypothesis was that exposure to light after birth could induce the change in expression of the skin-transcript in the skin. Because investigation of factors affecting expression of the skin-specific transcript could more easily be performed *in vitro*, primary culture experiments were carried out on cells isolated from C57Bl/6 mice from various developmental stages. The normal transcript is not UV inducible but the skin-specific transcript required investigation. No skin-specific transcripts have been reported for the other genes involved in NER but skin is a difficult tissue to work with due to the high levels of nucleases that degrade RNA (another reason why *in vitro* investigations of transcript induction are preferable) and it may simply be that expression patterns in skin have not received much attention. The work on the keratinocyte pools in the minigene transfections had shown an increase in upper transcript production in response to UV-C irradiation and the observation that the level and ratio of skin-transcript increased after birth suggested that the novel transcript could, indeed be UV inducible.

Isolation of cells from various developmental stages required different methods because of the difficulty associated with separation of the skin from the rest of the tissue in the early stages. This meant that E13-E15 isolated fibroblast cells were from dissaggregated whole embryos rather than dermis. In the later stages of development the dermal fibroblasts were isolated by removal of skin and separation of dermis from epidermis. In the primary culture analysis, northern blots were performed rather than RT-PCR because northern blots enable more quantitative comparison of expression levels.

In experiment 6.1 the *Ercc1* transcript expression patterns of the primary cultures were compared to that of the corresponding developmental stage and differences were observed. The developmental pattern observed in the *in vivo* samples followed that described earlier (Dr. Winter, unpublished data), with increasing levels of skin-specific transcript and ratio of skin-specific to normal transcript as development progressed. However, the expression pattern observed in the cultures isolated from the developmental stages differed and was more similar to developmental expression patterns observed later in development *in vivo*. It was as if culturing the cells

accelerated the change in skin-specific expression observed *in vivo*. What was more remarkable was that this happened in the absence of light as these cultures had been isolated and cultured with minimal light exposure and the small exposure during isolation had been to natural daylight through glass (rather than direct sources of UV such as fluorescent lighting). This either meant that this small exposure to visible light had induced the changes or that these changes occur in the absence of light and UV. What was clear was that a factor related to culture of the cells induced a change in skin-specific transcript expression.

The changes occurring in culture were related to proliferation in culture. In experiment 6.2, E13.5, E18.5 and P2 cells were isolated and cultured, with RNA sampling taking place during culture. The level of skin-specific transcript and the ratio of skin-specific to normal transcript increased with time in culture for cells from all three developmental stages (even though the later skin-specific transcript expression pattern was atypical for the corresponding developmental stage) and this occurred in the absence of light/UV. This indicated that the level of the transcript was affected by proliferation in culture. The possibility that growth factors in the serum could be involved is discussed later.

UV-B and UV-C were both tested as possible skin-specific transcript induction factors (experiment 6.3). As detailed in the introduction to chapter 6, UV-B and UV-C are both components of the solar spectrum and exposure to either results in DNA damage in the skin. UV-C is not naturally encountered as it is filtered out by the ozone layer while UV-B exposure varies according to a number of factors, for example, angle of the Sun in the sky). Both types of UV irradiation result in the same sort of damage but both were tested because UV-B is encountered in nature. For a set source height, a dose (Jm^{-2}) is more quickly supplied by UV-C than UV-B because UV-C has a shorter wavelength. Irradiation results in the same types of DNA damage with both UV-C and UV-B, primarily CPDs and (6-4)PPs. Interstrand crosslinks are very disruptive to cellular processes such as transcription and division as the DNA strands cannot be separated at the site of the interstrand cross-link. In the multiple irradiation experiments, dermal fibroblasts isolated from E18.5 dermis were cultured in darkness and repeatedly exposed to either UV-C, UV-B or visible

light. *In vivo*, cells from this developmental stage have a very low skin-specific transcript ratio and because the embryos are *in utero* they are not exposed to light. If cells for this experiment had been isolated from pups, the exposure to light following birth would have been a variable in the experiment. Because the skin-specific to normal transcript ratio and the level of transcript had been shown to increase in the absence of light in experiment 6.1, irradiated samples could only be compared to non-irradiated samples that had been cultured in darkness for the same length of time as the irradiated sample. The 'dark controls' were also the controls for time in culture changes. The experiment showed that there was only a very small difference between the skin-specific to normal transcript ratio and expression level when light and UV-B irradiated cultures were compared to dark controls and this was unlikely to be significant. However, there was a massive increase in the skin-specific to normal transcript ratio (U:L) in the UV-C irradiated sample when compared to the dark control. This suggested that expression of the skin-specific transcript could possibly be induced by UV-C irradiation. This result supports the earlier observation that skin-specific transcript production was increased by UV-C irradiation of keratinocyte pools transformed with *Ercc1* MG#20 (experiment 5.6). Because the CT repeat region is deleted from *Ercc1* MG#20, UV induction of the skin transcript could not involve interaction between transcription factors expressed/activated following UV irradiation and the CT repeat region. UV-induction of the skin-specific transcript would have to be via another part of the 5′ flanking sequence. The apparent induction of skin-specific transcript expression would require further experimentation before firm conclusions regarding the UV-C inducibility of the skin-specific transcript could be made. UV-C has a shorter wavelength than UV-B and DNA absorbs maximally within the wavelength range of UV-C. It is possible that the skin-specific transcript is induced by UV in a DNA damage-dependent manner. Equivalent doses of UV-C and UV-B are not absorbed equally by DNA so perhaps a higher dose of UV-B would have resulted in induction of skin-specific transcript expression. This possibility could be addressed in future work.

The non-light related changes in skin-specific transcript expression observed in cultured cells were also investigated further. The effect of serum starvation upon

Ercc1 expression in E18.5 murine primary culture was interesting (experiment 6.4). During serum starvation the level of both the skin-specific and normal transcripts dropped. However, following release from serum starvation there was a decrease in the normal transcript production and a large increase in skin-specific transcript production with the highest skin-specific to normal transcript ratio 6h post release and highest *Ercc1* production 24h post release. What was also interesting was that the skin-specific to normal transcript ratio and the sum *Ercc1* transcript level following release from serum starvation exceeded the pre-starvation level. In comparison, the skin-specific to normal transcript ratio of the non-starved control decreased during the course of the experiment while the sum *Ercc1* level increased. Serum starvation causes arrest of the cell cycle in the G1 stage and this response is due to the reduction in the supply of growth factors. Following release from serum starvation there are two categories of gene response; the first genes to be expressed following release are the 'early response' genes and these are mainly transcription factors required for subsequent expression of the 'late response' genes. The recovery pattern of *Ercc1* expression following serum starvation fits with the 'late response' pattern and it is therefore likely that the transcription factor of *Ercc1* is an 'early response' gene. If the level of transcription factor following recovery from serum starvation was higher than prior to serum starvation then that would be a method by which expression of the skin-specific transcript could be higher following release from serum starvation than prior to starvation. Serum may also explain the observation that *Ercc1* expression changes following isolation and culture, with expression of the skin transcript in cells isolated from developmental stages at which the skin-transcript would not normally be expressed. As long as the cells are *in vivo* the growth factors are being supplied in a developmentally regulated way. Once the cells are isolated and cultured in serum, it may be that the growth factors supplied cause production of transcription factors that would not normally be produced *in vivo* until later in development. The cells could then react by increasing production of the skin-specific *Ercc1* transcript. This response would be programmed into development and the pattern of skin-specific *Ercc1* transcript expression observed in the skin would then be tissue-specific rather than induced, although the two mechanisms need not be mutually exclusive.

It is possible that the CT repeat acts as a tissue-specific enhancer of skin-specific transcript production. Although the CT repeat is 3′ to the skin-specific transcript's TSS, CT repeats can act from downstream so the location of the CT repeat need not preclude this activity. This speculation is supported by the discovery that followed from experiment 6.5, the discovery being a size difference between the skin-specific transcript expressed by fibroblasts (derived from C57Bl/6) and keratinocytes (derived from BALB/c). Unfortunately this difference was identified very late in the research period, so the sequence work to investigate the difference in size was carried out by Dr. Winter (unpublished data). The size difference was found to be due to a strain difference between the mice with an extra 35 bps present in the CT repeat region of the *Ercc1* gene's 5′ flanking region in BALB/c mice.

Expression of upper transcript in different mouse strains: Strain specific differences in the size of the *Ercc1* upper transcript were discovered during the course of the experiments presented in this thesis. BALB/c mice have an extra 74 base pairs in their CT repeat region of their *Ercc1* gene when compared to C57Bl/6 mice. BALB/c mice also have an extra 183bp between nucleotides -1046 and -1047 of the 5' flanking sequence (5' to the CT repeat region) - see figure 6.7 for size comparison of *Ercc1* upper transcripts. C57Bl/6 mice and 129/Ola mice have the same size of CT repeat, while BALB/c mice have an extra 35 bps. The CD-1 CT repeat information is not available at this time. Figure 7.1 shows that the upper transcript expression level in the skin of mice from the mentioned strains increases with the level of skin-specific *Ercc1* transcript expressed, although quantification would be required before conclusions on proportionality can be drawn. In figure 7.1, the loading is equal. There are differences in the level of upper transcript expressed between the mouse strains. The level of upper transcript is lower in the mouse strains that have the smaller number of CT repeats. C57Bl/6 and 129/Ola have low levels of upper transcript expression and both have 35 fewer bp in their CT repeat regions that BALB/c; BALB/c mice have a high level of upper transcript expression. The number of CT repeats in the CD-1 mice has not been determined but this strain has a higher level of upper transcript expression than the BALB/c mice, which may

be due to strain variation in the 5' promoter region of the *Ercc1* gene. As removal of the CT repeat region seemed to cause a reduction in upper transcript expression in transformed keratinocytes (chapter 5, assuming RT-PCR product band intensity is proportional to transcript level), the observations relating to strain variation in CT repeat number and *Ercc1* upper transcript expression would appear to support the importance of the CT repeat region to expression of the upper transcript, assuming product band intensity is proportional to transcript level in the RT-PCR data.

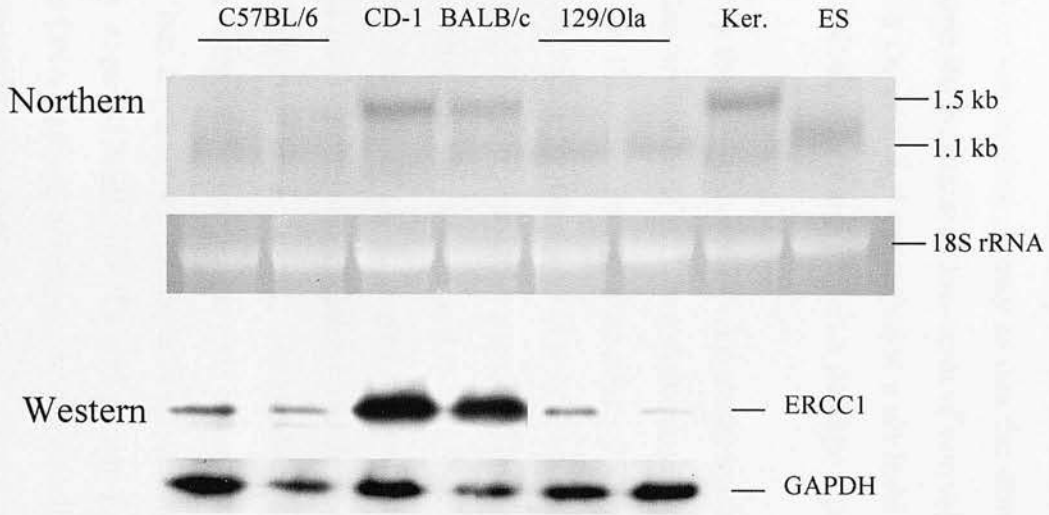
CT repeats result in formation of non-B-DNA structures and increasing the number of CT repeats has been shown to increase non-B-DNA structure formation and resulting transcription of a reporter gene (Xu and Goodridge, 1998). CT repeats are often found in promoters that lack classical promoter elements and are capable of initiating transcription (Weis and Reinberg, 1992 and 1997). Developmentally regulated and tissue specific transcription from genes with CT repeat regions has also been reported (Bevilacqua *et al.*, 2000). It would be interesting to investigate whether the skin-specific transcript expression could be induced by higher levels of EGF in growth medium via the CT repeat region. Growth factors, including EGF, have been shown to interact with a protein that binds to CT repeats (Wyse *et al.*, 2000) and CT repeats have been identified in genes that have more than one TSS, such as the chicken malic enzyme gene, where the $(CT)_7$ repeat binds single-stranded DNA-binding protein and is essential for promoter activity (Hodnett *et al.*, 1996; Xu and Goodridge, 1996). CT repeats regulate transcription of genes in a diverse range of organisms, $(CT)_n = (GA)_n$, including the *Drosophila* heat shock proteins *hsp26* and *hsp70* in which GAGA transcription factor (GAF) binds the site where it can cause promoter distortion (Lu *et al.*, 1993; Wilkins and Lis, 1998 and 1999). A novel murine GAGA box-binding factor, which is structurally related to the *Drosophila* GAGA factor has been reported. This factor was isolated by *in vivo* transcription factor titrations with double-stranded oligodeoxyribonucleotides and antibodies, perhaps a similar method could be used to isolate any factors interacting with the CT repeat region in the mouse *Ercc1* gene (Bevilacqua *et al*, 2000). The skin-specific expression pattern of the *Ercc1* skin-transcript could result from control of expression in a tissue-specific manner via the CT repeat region.

An obvious difference between the mouse strains mentioned is their coat colour. C57Bl/6 mice are black, 129/Ola are light yellow, BALB/c and CD-1mice are non-pigmented (albino). The two albino strains had higher expression of upper transcript than the pigmented strains. If there was greater biological demand for high skin-specific transcript expression in mice lacking protective dark pigmentation then this would be good evidence for the importance of the upper transcript in the skin to meeting biological demand for UV damage repair. Experiments performed subsequent to the end of my research involved crossing BALB/c and C57Bl/6 mice to segregate the *Ercc1* gene from those determining coat colour. Analysis of the F2 mice revealed that the C57Bl/6 or BALB/c *Ercc1* allele (or something tightly linked to it) determines the level of expression. Recombinant black mice containing the Balb/c *Ercc1* allele and recombinant white mice containing the C57Bl/6 *Ercc1* allele are being studied to determine the biological significance of the level of skin-specific *Ercc1* transcript expression on a non-pigmented background.

Figure 7.1 Level of expression of *Ercc1* skin-specific transcript is directly proportional to the level of *Ercc1* protein in skin.

This work was performed by Dr. Winter, Melton Lab. The upper panels show the northern analysis of the *Ercc1* expression in four mouse strains (C57BL/6, CD-1, BALB/c and 129/Ola) with keratinocyte and ES cell controls. The ethidium bromide staining of the 18S rRNA is shown as a loading control. The western analysis shows the corresponding level of Ercc1 protein in the skin and the GAPDH reprobe is the loading control.

With regard to a possible correlation between CT repeat region and skin-specific transcript expression, C57BL/6 mice and 129/Ola mice have the same size of CT repeat while BALB/c mice have an extra 35bps. The CD-1 mouse CT repeat information is not available at this time.


C57BL/6
CD-1
BALB/c
129/Ola
Ker.
ES
Northern
1.5 kb
1.1 kb
18S rRNA
Western
ERCC1
GAPDH

The possible UV-C induction of the skin-specific transcript expression, speculatively reported in this thesis, may involve a different method of skin-specific transcript expression control. The possible UV-C induction of skin-specific transcript expression in keratinocytes pools transformed with *Ercc1* MG#20, which lacks the CT repeat region, suggests the possibility of a CT repeat independent mechanism of UV-C induction. Such a method of regulation would allow the steady-state level of skin-specific transcript to be high and under control of e.g. the CT repeat region, while also allowing for induction of skin transcript expression at times of high UV-induced damage by a separate mechanism. It may be that the skin-specific transcript is required in the skin because the constitutive low-level of normal transcript is under the control of a transcription factor that is expressed at a set level and this cannot be altered to meet the higher biological demand for DNA damage repair in the skin.

Speculatively, a possible way that UV-induced DNA damage could be linked to skin-specific transcript production would be if p53 or a downstream transcription factor was a transcription factor for the skin-specific transcript. p53 is a multifunctional protein and is a transcription factor for over 100 genes. The protein is involved in regulation of apoptosis, the cell cycle and DNA repair. The work linking p53 to regulation of DNA damage repair is very recent and is reviewed in Adimoolam and Ford, 2003. Following exposure to DNA damaging agents such as UV, the p53 protein is stabilised via post translational modifications and downstream target genes containing the p53 response element are activated. The products of these genes may, in turn, act as transcription factors for other genes. p53 can also interact via protein-protein interactions. p53 has been shown to increase expression of a number of genes involved in the early stages of NER. mRNA and protein products of the *XPC* gene increase in a p53 and DNA damage-dependent manner (Adimoolam and Ford, 2002). This ratio of skin-specific transcript to normal transcript was dramatically increased by UV-C irradiation in experiment 6.3, so perhaps the same mechanism of induction could be taking place as with the *XPC* gene. The *XPC* gene promoter contains a p53 response element and the *Ercc1* gene's 5′ flanking region could also be searched for such an element. p53 directly binds to the NER proteins XPB and XPD (subunits of TFIIH), where it has been reported to inhibit helicase activity

(Wang *et al.*, 1995). The product of the *BRCA1* gene has also been shown to increase expression XPC in response to DNA damage and is, therefore, involved in the DNA damage recognition step of NER independently of p53 (Hartman and Ford, 2002; Adimoolam and Ford, 2003). The responsiveness of skin-transcript expression to this gene product could also be investigated. XPC is required for GGR but is not essential for TCR (Venema *et al*, 1990 and1991) and cells deficient in p53 are deficient in GGR. XPC is essential for removal of CPDs and (6-4) PPs *in vivo* (Emmert *et al.*, 2000). Within minutes of UV irradiation, XPC binds to UV-induced CPDs, along with another protein expressed in response to p53 levels called p48. The basal level of these proteins controls the initial rate of GGR of CPDs. It has been suggested that the function of the UV-induction (via p53) of these proteins is to replenish the level of these repair factors following proteosome mediated degradation to ensure that the levels are sufficient for lesion repair. It has also been suggested that for these genes, the basal level of expression is important for the repair of the (6-4) PPs while the inducible levels are more important for the repair of CPDs. It may be that the high level of skin-specific transcript in the skin (perhaps caused by a tissue-specific response to the CT repeat region) is important for the removal of (6-4) PPs while the possible UV-induction of the skin transcript (perhaps by p53 mediated expression in response to UV-induced DNA damage) is important for removal of CPDs in the same manner as with the p53 mediated XPC and p48 response mentioned.

The skin-specific *Ercc1* transcript expression appeared to be induced by UV-C irradiation and also by proliferation in culture in experiments (6.1-6.4). UV-C induction of the DNA ligase I gene, which is involved in NER, has also been reported and expression of this gene is also induced by cellular proliferation. This work was performed in human fibroblasts and p53 was proposed as a possible factor in the UV-induction (Montecucco *et al.*, 1992 and1995). Further work on the UV-induction of the *Ercc1* skin-transcript need to be carried out. Because of the possibility that immortalised cell lines have distrupted p53 response/production and the possibility that p53 could be involved in skin-specific transcript expression in response to DNA damage, the p53 status of cells used for the experiments should

perhaps be ascertained prior to the UV-induction experiments with the *Ercc1* skin-specific transcript. If p53 was required for induction of the skin-specific transcript following UV irradiation, DNA damage could accumulate more quickly in skin cells that were becoming cancerous and had lost p53 activity, due to lack of p53 resulting in loss of the UV-induction response of the skin-specific transcript causing lower *Ercc1* protein levels and less DNA damage repair. A link between proliferation in tumour cells and skin-specific transcript production could be important biologically as a means to reduce the rate of damage accumulation in such cells; DNA ligase I is expressed at higher levels in human tumours than in benign normal tissues and it has been suggested to be more important in proliferating tissue than in resting cells (Sun *et al.*, 2001). The levels of skin-specific transcript and the ratio of skin-specific to normal transcript increased in response to proliferation in culture and release from serum starvation (exp. 6.1, 6.2 and 6.4) so the presence of the skin-specific transcript may also be more important in proliferating tissues than resting cells.

Interestingly, considering the possible induction of the skin transcript in response to UV-C irradiation but not UV-B in experiment 6.3 and, a wavelength specific effect has been observed in the p53 mediated-NER. Two similar experiments in which CPDs were introduced to both the transcribed and non-transcribed strand of the p53 gene demonstrated that when the lesions were introduced using UV-B, GGR and TCR was enhanced by the addition of functional p53, suggesting that p53 was required for both GGR and TCR. However, when UV-C was used the results suggested that p53 was only involved in TCR (Mathonnet *et al.*, 2003). This highlights a difference in NER control by p53 in response to UV-B or UV-C induced DNA damage. In the UV-irradiation experiment 6.3, the skin-transcript to normal transcript ratio rose dramatically in response to the UV-C irradiation but not the UV-B irradiation. If the skin-specific transcript expression is induced by p53 then this might be why there was a difference between the skin-specific transcript expression response to UV-C and UV-B induced DNA damage. The choice of UV source will need to be considered in future work on *Ercc1* if this is the case since induction of the skin-specific transcript may be wavelength-dependent. Work on skin-transcript induction *in vivo* in our laboratory used UV-B and the transcript was not induced.

However, because UV-B is encountered in nature while UV-C is filtered out of the solar spectrum before reaching ground level, UV-B is the more biologically relevant DNA damage source. The possibility of induction of the skin-transcript by UV-C but not UV-B means that the choice of UV source for future work on the skin-specific transcript and more generally in terms of modelling cancer development in mice will require consideration. If the novel-transcript is induced by UV-C and not UV-B this would be once of very few examples of a UV-C/UV-B specific response reported. This is interesting because these sources are presumed to cause similar biological effects in most studies of this kind.

To revisit the original aims of the research presented in this thesis, the minigene work was performed to identify elements of the 5′ flanking region of the *Ercc1* gene that could be responsible for the observed skin-specific *Ercc1* expression pattern. *Ercc1* MG#13 has been shown to contain the necessary sequences for skin-specific transcript expression. Deletion of the CT repeat region resulted in alteration of skin-specific transcript expression. The complete absence of skin-specific transcript following transfection with MG#20 in contrast to expression of this transcript in cells transfected with MG#13 was reported. However, most differences were in terms of the intensity of product band in RT-PCR (generally a reduction in intensity) in cells that normally express the skin-specific transcript. As explained previously, the possible link between the level of transcript and the band intensity is speculative in the absence of quantitative RT-PCR. The CT repeat region may be part of the skin-specific transcript promoter or an enhancer region for the skin-specific transcript but further work is required, especially quantitative RT-PCR, before any conclusions may be drawn upon the importance of this region to expression of the skin-specific transcript. In the absence of such work, the observations upon the effects of deletion of this region are intriguing but remain speculations only. The response of the skin-specific transcript to proliferation in culture and UV-C irradiation led to speculation upon the possibility that the transcription factor p53 could be involved in UV-induced skin-specific transcript expression. This transcription factor is known to be involved in expression of a number of the genes involved in NER, although a role in

Ercc1 expression has not been reported and investigation of the skin-specific transcript expression has not been reported.

Wider approaches that could be used to study the importance of the novel *Ercc1* transcript and indeed the normal *Ercc1* transcript include tissue-specific complete knock-out or disruption of the *Ercc1* gene in the skin, knock-out of one of the transcripts in the skin or tissue-specific expression of one or both of the transcripts *in vivo* and expression of only the skin-specific transcript *in vitro*. There are a number of approaches that can be used to disrupt the *Ercc1* gene or introduce it *in vitro* and *in vivo*. Genes can be disrupted or knocked out in a number of different ways and the choice of technique depends upon the suitability of the delivery vector and the purpose of the modification. The role of the *Ercc1* gene has been investigated in a number of ways in the Melton laboratory prior to the research presented in this thesis and further progress has been made into the investigation of of *Ercc1* expression in the skin during and since the end of the research reported here, so these discussions are not hypothetical.

Techniques for introduction of genetic material into cultured cells include the use of viruses and minigenes. Retrovirus-mediated gene transfer of functional *Ercc1* into *Ercc1*-null cultured cells was attempted prior to my arrival in the Melton lab. A recombinant retrovirus containing the *Ercc1* coding sequence was manufactured using the murine leukaemia virus derived, MFG viral vector (to ensure specificity of the vector for the murine system) and the viral packaging cell line ψCRE. The *in vitro* experiments in which *Ercc1*-deficient fibroblasts were exposed to the retroviruses and then tested for rescue of UV hypersensitivity were performed as a precursor to *in vivo* gene therapy with liver specificity for *Ercc1* null mice to extend their lifespan and allow long-term study of the effect of *Ercc1* knockout in other tissues. Retroviruses are useful as viral genomes which integrate into the host cell should be stably expressed for the lifetime of the infected cell. Following retroviral-mediated gene transfer there was partial correction of the UV-hypersensitivity of the cells but it was decided that a better technique for introduction of a gene was use of a minigene system. Because skin does divide rapidly the skin is a suitable tissue for

infection by a retrovirus designed to have tropism for this tissue (cell division is a prerequisite for infection). However, this would be a technically demanding, expensive course to take and would involve overcoming many obstacles before progress would be likely. In vitro investigation of the role of the skin-specific transcript could be performed in the same manner as the minigene work presented if the sequences responsible for expression of the transcripts could be accurately identified and deleted. Minigenes containing the *Ercc1* DNA and the promoter for only one of the transcripts could be used to transform any of the cells used in these studies to investigate the ability of the skin-specific transcript, in the absence of the normal transcript, to correct UV hypersensitivity of *Ercc1*-null cells. As the novel transcript is not normally found alone *in vivo*, there would be little to gain from developing mice that expressed only the novel transcript in terms of investigation of the biological significance of the novel transcript.

If the skin-specific *Ercc1* transcript is biologically significant then the most obvious way of detecting this is by loss of function via targetted deletion. However, depending upon the question being investigated the normal transcript's function may or may not be required and this has implications for the knock-out method. The *Ercc1* complete knockout mice are discussed in details in Chapter 1; due to their reduced lifespan and the complicating effects of the loss of *Ercc1* in other tissues, their usefulness for specific investigation of the skin-specific transcript function is limited. A better way of investigation this transcript's role is a tissue specific knockout. Another consideration in production of a tissue-specific knockout is whether temporal control of the deletion is desirable (conditional knock-out) as such a deletion enables study of the consequences of deletion at specific developmental stages. Another way to investigate the biological significance of removing one or both of the transcripts would be to use a technique called RNA interference (RNAi).
A relatively new technique for investigation of gene function is RNA interference (RNAi). This technique results in 'knock-down' of genes - the genes are expressed but the mRNA is bound by synthetic short interfering RNAs that specifically target the mRNA for degradation. In this way an actively transcribed gene does not result in translation of a protein. This is a naturally occurring mechanism for gene silencing

in a number of species, e.g. in *Caenorhabditis elegans* dsRNA can silence homologous genes. The precursor long, dsRNA is synthesised and an enzyme called Dicer cleaves the dsRNA into 21-23 nucleotide long duplexes called small interfering RNAs (siRNA). When these siRNAs are bound by a multiprotein complex called RNA induced silencing complex (RISC) the siRNA unwinds and binds complementary mRNA. This targets the bound mRNA for degradation by RISC via the mRNA degredation pathway (Cogini and Macino, 2000; Hammond *et al*, 2001; McManus *et al*, 2002).

This naturally occurring gene silencing method can be adapted to knock down genes in mammals if synthetic siRNA is designed and transfected into cells directly. Introduction of the precurser long, double stranded RNA results in degradation via the interferon response in the cytoplasm. Knock down of genes in mice has been achieved by expressing small hairpin RNA (shRNA) from a polymerase III promoter.

Knock downs of a gene in a tissue-specific manner depends upon use of a tissue-specific promoter to drive expression of long dsRNA. Avoidance of the interferon response has been achieved by use of a vector called pDECAP. This vector expresses the dsRNA but removes the 5' and 3' caps, restricting the dsRNA to the nucleus. This dsRNA is cleaved inside the nucleus and the siRNAs are transported into the cytoplasm to knock down homologous mRNA. If a tissue specific promoter is used in the plasmid DNA vector then the RNAi can occur in the tissue of interest, resulting in tissue specific knock-down of a gene e.g. knock-down of *Ercc1* in the skin. However, as the mechanism is though to occur via the 3′ end of the short interfering RNA and the difference between the two *Ercc1* transcripts lies in the 5' end, it may not be possible to knock-down the novel or normal transcript independently. The RNAi technique may be useful for complete knock-down of the *Ercc1* gene in the skin, although this has already been achieved using another method, the Cre/*lox* system.

Because of the severe phenotype associated with the *Ercc1* knockout mice, their usefulness for investigation of the skin-specific *Ercc1* transcript is limited. Skin specific deletion of the novel transcript with the option of temporal control of expression would enable specific investigation of the biological significance of the transcript in the skin. Because both the novel and normal *Ercc1* transcripts utilise the same coding region, knockout of only the novel transcript requires identification of sequences particular to the novel transcript so that their deletion does not interfere with normal transcript expression and function.

The methods used for generation of the complete *Ercc1* knockout are detailed in Chapter 1; the Melton lab. *Ercc1* knockout was generated using the embryonic stem cell method and involved disruption of the *Ercc1* gene using a neo-cassette which was targeted by homologous recombination. This resulted in loss of *Ercc1* function in all tissues. For tissue specific deletion of *Ercc1*, a conditional knockout method would need to be utilised. In Cre/*lox,* the *cre* gene (short for cyclisation recombination) encodes a site specific DNA recombinase which has the property of recombining DNA located between specific sites in the DNA molecule, *loxP* sites (locus of crossover over P1). *loxP* sites are 34bp long and are not found naturally in animals or plants as they originate from the P1 bacterial phage. *LoxP* sites can be artificially introduced to enable precise excision of DNA using *cre* under control of a tissue-specific promoter. DNA excised, e.g. part of the novel *Ercc1* transcript's promoter, is degraded and *loxP* sites flanking the excised region are ligated (Ghosh and Van Duyne, 2002; Kuhn and Torres, 2002; Le and Sauer, 2000; Sauer, 1998). Skin-specific hairless *Ercc1* knockout mice have been produced in the Melton lab using the Cre/*lox* system. Male mice carrying the *cre* gene with the genotype *Ercc1* KTΔ/+ crossed with an *Ercc1* flox (*Ercc1* flanked by loxP stites) result in progeny which include pups of the genotype Δ/flox + *cre*. The *cre* is under control of a CK5 (a keratin) promoter so this skin specific promoter results in expression of the *cre* in the skin, resulting in excision of the floxed *Ercc1* gene in a tissue specific-manner. The breeding scheme is complicated by the fact that the *cre* must be inherited from the father because of brief expression of *cre* in the oocyte of *cre*+ females (Ramirez *et al.*, 2004). Depending upon the promoter used to drive the *cre* gene, tissue

specific or temporal control of *cre* expression can be achieved. The Cre/*lox* system involves deletion in the genome and for specific deletion of promoter sequences required for expression of the skin specific transcript there would be the possibility that expression of the normal transcript could also be disrupted, depending upon the exclusivity of the promoter's regions in the *Ercc1* gene. There are, therefore, numerous ways in which the role of *Ercc1* transcripts in the skin could be/are being investigated further.

The research investigating the role of the skin-specific transcript presented in this thesis demonstrated that the level of transcript and the ratio of skin-specific to normal transcript rose in the absence of light in response to proliferation in culture and also to return of serum after serum starvation (a cell cycle dependent response). There were also results that suggested expression of the transcript could be induced by UV-C irradiation. The earlier question posed regarding the requirement for the skin-specific transcript in the skin has been addressed. If up-regulation of the normal transcript could lead to increased protein levels in conditions of high biological demand for DNA repair then the skin-specific transcript would not be required. The normal transcript is expressed constitutively in the body at low levels and is required for repair of endogenous DNA damage in tissues that are not exposed to light as well as in the skin. It may be that the control of expression of the normal transcript is incompatible with increased expression in the skin as a steady-state level of this transcript is achieved by a static expression level of the normal transcript's transcription factor. However, the increased biological demand for repair of UV-induced DNA damage in the skin, primarily CPDs and (6-4)PPs, may have resulted in evolutionary selection for a different pattern of expression in this tissue. The development of the high expression level and ratio to normal transcript of the skin-specific transcript occurs in the absence of light and may be due to enhanced expression of the skin-specific transcript via a tissue-specific and developmental response regulated by interaction with the CT repeat region located near the TSS of the skin-specific transcript. This mechanism of control could enable a high basal level of skin-specific transcript resulting in a high level of *Ercc1* protein for removal of UV (6-4)PPs. The possibility that UV-induction could increase the expression

level has speculatively been linked to the requirement for rapid removal of CPDs following UV irradiation. Because these lesions are very disruptive to cellular processes and non-repair leads to accumulation of DNA damage and possibly carcinogenesis, the high level of skin-specific transcript (whether by means of tissue-specific expression pattern and/or UV-inducible expression) is likely to confer a considerable advantage to the skin in removal of UV-induced DNA damage and avoidance of carcinogenesis. Because of the possibility of skin cancer development, especially melanoma, in response to lack of repair of UV-induced DNA damage in the skin it is likely that there has been strong evolutionary selection pressure on the expression of *Ercc1* in this tissue and the skin-specific pattern of expression has evolved as a result.

References

Aboussekhra, A., Biggerstaff, M., Shivji, M.K., Vilpo, J.A., Moncollin, V., Podust,V.N., Protic, M., Hubscher, U., Egly, J.M. and Wood, R.D. (1995) Mammalian DNA nucleotide excision repair reconstituted with purified protein components. *Cell.* **80** (6), 859-568.

Adimoolam, S. and Ford, J.M. (2003) p53 and regulation of DNA damage recognition during nucleotide excision repair. *DNA Repair* **2** 947-954.

Adimoolam, S. and Ford, J.M. (2002) p53 and DNA damage-inducible expression of the xeroderma pigmentosum group C gene. *Proc.Natl.Acad.Sci.U.S.A.* **99** 12985-12990.

Araujo, S.J. and Wood, R.D. (1999) Protein complexes in nucleotide excision repair. *Mutat. Res.* **435** 23-33.

Bardwell, A.J., Bardwell, L., Tomkinson, A.E. and Friedberg, E.C. (1994) Specific cleavage of model recombination and repair intermediates by the yeast Rad1-Rad10 DNA endonuclease. *Science* **265** (5181), 2082-2085.

Barre, F.X., Asseline, U. and Harel-Bellan, A. (1999) Asymmetric recognition of psoralen interstrand crosslinks by the nucleotide excision repair and the error-prone repair pathways. *J. Mol. Biol.* **286** (5), 1379-1387.

Batty, D., Rapic' -Otrin, V., Levine, A.S. and Wood, R.D. (2000) Stable binding of human XPC complex to irradiated DNA confers strong discrimination for damaged strand. *J.Mol.Biol.* **2000** 275-290.

Baumann, P. and West, S.C. (1988) DNA end-joining catalysed by human cell-free extracts. *Proc.Nat.Acad.Sci.U.S.A.* **95** 14066-14070.

Bevilacqua, A., Fiorenza, M.T. and Mangia, F. (2000) A developmentally regulated GAGA box-binding factor and Sp1 are required for transcription of the hsp70.1 gene at the onset of mouse zygotic genome activation. *Development.* **127** (7), 1541-1551.

Bohr, V.A., Smith, C.A., Okumoto, D.S. and Hanawalt, P.C. (1985) DNA repair in an active gene: removal of pyrimidine dimers from the DHFR gene of CHO cells is much more efficient than in the genome overall. *Cell* **40** 359-369.

Bridges, B.A. and von Wright, A. (1981) Influence of mutations at the rep gene on survival of Escherichia coli following ultraviolet light irradiation or 8-methoxypsoralen photosensitization: evidence for a recA+ rep+-dependent pathway for repair of DNA crosslinks. *Mutat. Res.* **82** (2), 229-238.

Chipchase, M.D., O'Neill, M. and Melton, D.W. (2003) Characterization of premature liver polyploidy in DNA repair (Ercc1)-deficient mice. *Hepatology* **38** (4), 958-966.

Chomczynski, P and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidium thiocyanate-phenol-chloroform extraction. *Anal.Biochem.* **162** (1) 159-159.

Cogoni, C., Macino, G. (2000) Post-transcriptional gene silencing across kingdoms. *Genes Dev.* **10** 638-643

Cook, J.S. (1970) Photoreactivation in animal cells. *Photophysiology* **5** 191-233.

Cui, X., Brenneman, M., Meyne, J., Oshimura, M., Goodwin, E.H. and Chen, D.J. (1999) The XRCC2 and XRCC3 repair genes are required for chromosomal stability in mammalian cells. *Mutat.Res.* **434** (75), 88.

Dagert, M. and Ehrlich, S.D. (1974) Prolonged incubation in calcium phosphate improves competence of *Escherichia coli* cells. *Gene* **6** 23-28.

de Boer, J., Donker, I., de Wit, J., Hoeijmakers, J.H. and Weeda, G. (1998) Disruption of the mouse xeroderma pigmentosum group D DNA repair/basal transcription gene results in preimplantation lethality. *Cancer Res.* **58** (1), 89-94.

de Laat, W.L., Appeldoorn, E., Sugasawa, K., Weterings, E, Jaspers, N.G. and Hoeikmakers, J.H. (1998) DNA binding polarity of human replication protein A positions nucleases in nucleotide excision repair. *Genes & Dev.* **12** 2598-2609.

de Silva, I.U., McHugh, P.J., Clingen, P.H. and Hartley, J.A. (2000) Defining the roles of nucleotide excision repair and recombination in the repair of DNA interstrand cross-links in mammalian cells. *Mol. Cell. Biol.* **20** (21), 7980-7990.

de Vries, A., van Oostrom, C.T., Hofhuis, F.M., Dortant, P.M., Berg, R.J., de Gruijl, F.R.,Wester, P.W., van Kreijl, C.F., Capel, P.J., van Steeg, H. and Verbeek, S.J. (1995) Increased susceptibility to ultraviolet-B and carcinogens of mice lacking the DNA excision repair gene XPA. *Nature* **377** (6545), 169-173.

de Vries, A., van Oostrom, C.T., Dortant, P.M., Beems, R.B., van Kreijl, C.F., Capel, P.J. and van Steeg, H. (1997) Spontaneous liver tumors and benzo[a]pyrene-induced lymphomas in XPA-deficient mice. *Mol.Carcinog.* **19** (1), 46-53.

Demple, B. and Harrison, L. (1994) Repair of oxidative damage to DNA: enzymology and biology. *Annu.Rev.Biochem.* **63** 915-948.

Dianov, G. and Lindahl, T. (1994) Reconstitution of the DNA base excision-repair pathway. *Curr.Biol.* **4** 1069-1076.

Dip, R., Camenisch, U. and Naegeli, H. (2004) Mechanisms of DNA damage recognition and strand discrimination in human nucleotide excision repair. *DNA Repair* **3** (11), 1409-1423.

Drapkin,R. and Reinberg, D. (1994) The multifunctional TFIIH complex and transcriptional control. *Biochem.Sci.* **19** 504-508.

Emmert, S., Kobayashi, N., Khan, S.G. and Kraemer, K.H. (2000) The xeroderma pigmentosum group C gene leads to selective repair of cyclobutane pyrimidine dimers rather than 6-4 photoproducts. *Proc.Nat.Acad.Sci.U.S.A.* **97** 2151-2156.

Evans, E., Moggs, J.R., Hwang, J.R., Egly, J.M. and Wood, R.D. (1997) Mechanism of open complex and dual incision formation by human nucleotide excision repair factors. *EMBO* **16** 6559-6573.

Feinberg, A.P. and Vogelstein, B. (1983) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Anal.Biochem.* **132** 6-13.

Fortini, P., Pascucci, B., Parlanti, E., D'Errico, M., Simonelli, V. and Dogliotti, E. (2003) The base excision repair: mechanisms and its relevance for cancer susceptibility. *Biochimie.* **85** (11), 1053-1071.

Friedberg, E.C., Walker, G.C., Siede, W. *DNA repair and mutagenesis.* Washington DC. ASM Press; 1995

Friedberg, E.C. (1996) Relationships between DNA repair and transcription. *Annu.Rev.Biochem.* **65** 15-42.

Friedberg, E.C. and Meira, L.B. (2003) Database of mouse strains carrying targeted mutations in genes affecting biological responses to DNA damage. *DNA Repair* **2** (5), 501-530.

Friedberg, E.C. (2004) The discovery that xeroderma pigmentosum (XP) results from defective nucleotide excision repair. *DNA Repair* **3** (2), 13-195.

Frosina, G., Fortini, P., Rossi,O., Carrozzino, F., Raspaglio, G., Cox, L.S., Lane, D.P., Abbondandolo, A. and Dogliotti, E. (1996) Two pathways for base excision repair in mammalian cells. *J.Biol.Chem.* **271** 9573-9578.

Fujiwara, Y., Masutani, C., Mizukoshi, T., Kondo, J., Hanaoka, F. and Iwai, S. (1999) Characterization of DNA recognition by the human UV-damaged DNA-binding protein. *J.Biol.Chem.* **274** 20027-20033.

Ghosh, K., Van Duyne, G.D. (2002). Cre-loxP Biochemistry. *Methods* **28** (3) 374-83.

Greenberg, R.B., Alberti, M., Hearst, J.E., Chua, M.A. and Saffran, W.A. (2001) Recombinational and mutagenic repair of psoralen interstrand cross-links in Saccharomyces cerevisiae. *J. Biol Chem.* **276** (34), 31551-31560.

Gunz, D., Hess, M.T. and Naegeli, H. (1996) Recognition of DNA adducts by human nucleotide excision repair. *J.Biol.Chem.* **271** (41), 25089-25098.

Hammond, S.M., Caudy, A.A., Hannon, G.J. (2001) Post-transcriptional Gene Silencing by Double-stranded RNA. *Nature Rev Gen.* **2** 110-119.

Hanahan, D. (1983) Studies on transformation of *Escherichia coli* with plasmids. *J.Mol.Biol.* **166** 557-589.

Hanawalt, P.C., Cooper, P.K., Ganesan, A.K. and Smith, C.A. (1979) DNA repair in bacteria and mammalian cells. *Ann.Rev.Biochem.* **48** 783-836.

Harm, W. (1970) Analysis of photoenzymatic repair of UV lesions in DNA by single light flashes. V. Determination of the reaction-rate constants in *E. coli* cells. *Mutat Res.* **10** (4), 277-290.

Hartman, A.R. and Ford, J.M. (2002) BRCA1 induces DNA damage recognition factors and enhances nucleotide excision repair. *Nat.Genet.* **32** 180-184.

Hawk, J. and McGregor, J. *Skin and Sunlight.* London, Dorling Kindersley Ltd; 2000

Hayashi, T., Takao, M., Tanaka, K. and Yasui, A. (1998) *ERCC1* mutations in UV-sensitive Chinese hamster ovary (CHO) cell lines. *Mutat.Res.* **407** 269-276.

Hodnett, D.W., Fantozzi, D.A., Thurmond, D.C., Klautky, S.A., MacPhee, K.G., Estrem, S.T., Xu, G. and Goodridge, A.G. (1996) The chicken malic enzyme gene: structural organization and identification of triiodothyronine response elements in the 5'-flanking DNA. *Arch Biochem. Biophys.* **334** (2), 309-324.

Hoeijmakers, J.H.J., van Duin, M., Westerveld, A., Yasui, A. and Bootsma, D. (1986) Identification of DNA repair genes in the human genome. *Cold Spring Harb.Symp.Quant.Biol.* **51** (1), 91-101.

Hooper, M., Hardy, K., Handyside, A., Hunter, S. and Monk, M. (1987) HPRT-deficient (Lesch-Nyhan) mouse embryos derived from germline colonization by cultured cells. *Nature.* **326** (6110), 292-295.

Hoy, C.A., Thompson, L.H., Mooney, C.L. and Salazar, E.P. (1985) Defective DNA cross-link removal in Chinese Hamster cell mutants hypersensitive to bifunctional alkylating agents. *Cancer Res.* **45** (1737), 1743.

Huang, J.C., Svoboda, D.L., Reardon, J.T. and Sancar, A. (1992) Human nucleotide excision nuclease removes thymine dimers from DNA by incising the 22nd phosphodiester bond 5' and the 6th phosphodiester bond 3' to the photodimer. *Proc.Nat.Acad.Sci.U.S.A.* **89** 3664-3668.

Husain, I., Griffith, J. and Sancar, A. (1988) Thymine dimers bend DNA. *Proc.Natl.Acad.Sci.U.S.A.* **85** 2558-2562.

Hwang, B.J. and Chu, G. (1993) Purification and characterization of a protein that binds to damaged DNA. *Biochemistry* **32** (1657), 1666.

Ivanov, E.L. and Haber, J.E. (1995) RAD1 and RAD10, but not other excision repair genes, are required for double-strand break-induced recombination in

Saccharomyces cerevisiae. Molecular and Cellular Biology **15** (4), 2245-2251.

Ivanov, E.L., Sugawara, N., Fishman-Lobell, J. and Haber, J.E. (1996) Genetic requirements for the single-strand annealing pathway of double-strand break repair in *Saccharomyces cerevisiae. Genetics* **142** (3), 693-704.

Jeggo, P.A. (1998) Identification of genes involved in repair of DNA double-strand breaks in mammalian cells. *Radiat. Res.* **150** (5 Suppl), 80-91.

King, R.W., Jackson, P.K. and Kirschner, M.W. (1994) Mitosis in transition. *Cell* **79** (4), 563-571.

Klugland, A. and Lindahl, T. (1997) Second pathway for completion of human DNA base excision repair: reconstitution with purified proteins and requirement for DNase IV (FEN1). *EMBO J.* **16** (3341), 3348.

Kolodner, R.D. and Marsischky, G.T. (1999) Eukaryotic DNA mismatch repair. *Curr.Opin.Genet.Dev.* **9** 89-96.

Kuhn, R., Torres, R.M. (2002). Cre/loxP Recombination System and Gene Targeting. *Methods Mol. Biol.* **180** 175-204.

Kunz, B.A., Straffon, A.F. and Vonarx, E.J. (2000) DNA damage-induced mutation: tolerance via translesion synthesis. *Mutat.Res.* **451** ((1-2)), 169-185.

Kuraoka, I., Kobertz, W.R., Ariza, R.R., Biggerstaff, M., Essigmann, J.M. and Wood, R.D. (2000) Repair of an interstrand DNA cross-link initiated by ERCC1-XPF repair/recombination nuclease. *J.Biol.Chem.* **275** (34), 26632-26636.

Le, Y., Sauer, B. (2000). Conditional Gene Knockout using cre Recombinase. Methods Mol. Biol. **136** 477-85.

Le Page, F., Kwoh, E.E., Avrutskaya, A., Gentil, A., Leadon, S.A., Sarasin, A. and Cooper, P.K. (2000) Transcription-Coupled Repair of 8-oxoGuanine Requirement for XPG, TFIIH, and CSB and Implications for Cockayne Syndrome. *Cell* **101** (2), 159-171.

Lehmann, A.R. (2000) Replication of UV-damaged DNA: new insights into links between DNA polymerases, mutagens and human disease. *Gene* **253** 1-12.

Lesch, M. and Nyhan, W.L. (1964) A familial disorder of uric acid metabolism and central nervous system function. *Am.J.Med.* **36** 561-570.

Lewis, L.K. and Resnick, M.A. (2000) Tying up loose ends: nonhomologous end-joining in *Saccharomyces cerevisiae*. *Mutat. Res.* **451** ((1-2)), 71-89.

Li, Y.F., Kim, S.T. and Sancar, A. (1993) Evidence for lack of DNA photoreactivating enzyme in humans. *Proc. Natl. Acad. Sci. U. S. A.* **90** (10), 4389-4393.

Lieber, M.R. (1999) The biochemistry and biological significance of nonhomologous DNA end joining: an essential repair process in multicellular eukaryotes. *Proc.Nat.Acad.Sci.U.S.A.* **4** (2), 77-85.

Lindahl, T. (1995) Recognition and processing of damaged DNA. *J.Cell.Sci.* **19** (Suppl.), 73-77.

Lindahl, T., Karran., P. and Wood, R.D. (1997) DNA excision pathways. *Curr.Opin.Genetics Dev.* **7** 158-169.

Liu, N., Lamerdin, J.E., Tebbs, R.S., Schild, D., Tucker, J.D., Shen, M.R., Brookman, K.W., Siciliano, M.J., Walter, C.A., Fan, W., Narayana, L.S., Zhou, Z.Q., Adamson, A.W., Sorensen, K.J., Chen, D.J., Jones, N.J. and Thompson, L.H. (1998) XRCC2 and XRCC3, new human Rad51-family members, promote chromosome stability and protect against DNA cross-links and other damages. *Mol.Cell.* **1** (6), 783-793.

Lu, Q., Wallrath, L.L., Granok, H. and Elgin, S.C.R. (1993) (CT)n.(GA)n repeats and heat shock elements have distinct roles in chromatin structure and transcriptional activation of the Drosophila *hsp26* gene. *J.Mol.Biol.* **13** 2802-2814.

Ma, Y., Lu, H., Tippin, B., Goodman, M.F., Shimazaki, N., Koiwai, O., Hsieh, C.L.,Schwarz,K. and Lieber,M.R. (2004) A biochemically defined system for mammalian nonhomologous DNA end joining. *Mol. Cell.* **16** (5), 701-713.

Mandel, M. and Higa, A. (1970) Calcium-dependent bacteriophage DNA infection. *Mol.Biol.* **53** 159-162.

Masutani, C., Kusumoto, R., Yamada, A., Dohmae, N., Yokoi,M., Yuasa, M., Araki, M., Iwai, S., Takio, K. and Hanaoka, F. (1999) The XPV (xeroderma pigmentosum variant) gene encodes human DNA polymerase eta. *Nature* **399** (6737), 700-704.

Mathonnet, G., Leger, C., Desnoyers, J., Drouin, R., Therrien, J.P. and Drobetsky, E.A. (2003) UV wavelength-dependent regulation of transcription-coupled nucleotide excision repair in p53-deficient human cells. *Proc. Natl. Acad. Sci. U. S. A.* **100** (12), 7219-7224.

McManus, M.T., Petersen. C.P., Haines, B.B., Chen, J., Sharp, P.A. (2002) Gene silencing using micro-RNA designed hairpins. *RNA.* **8**:842-850.

McWhir, J., Selfridge, J., Harrison, D.J., Squires, S. and Melton, D.W. (1993) Mice with DNA repair gene (*ERCC-1*) deficiency have elevated levels of p53, liver nuclear abnormalities and die before weaning. *Nat. Genet.* **5** (3), 217-224.

Mellon, I., Rajpal, D.K., Koi, M., Boland, C.R., Champe, G.N. (1996) Transcription-coupled repair deficiency and mutations in human mismatch repair genes. *Science* **272** (5261), 557-560.

Melton, D. W., Ketchen, A.M., Nunez, F., Bonatti-Abbondandolo, S., Abbondandolo, A., Squires, S. and Johnson, R.T. (1998) Cells from *ERCC1* -

deficient mice show increased genome instability and a reduced frequency of S-phase-dependent illegitimate chromosome exchange but a normal frequency of homologous recombination. *J.Cell Sci.* **111** 394-404.

Memisoglu, A. and Samson, L. (2000) Base excision repair in yeast and mammals. *Mutat.Res.* **451** (1-2), 39-51.

Moggs, J. G., Szymkowski, D. E., Yamada, M., Karran, P., and Wood,R.D. (1997) Differential human nucleotide excision repair of paired and mispaired cisplatin-DNA adducts. *Nucleic Acids Res.* **25** (480), 491.

Montecucco, A., Biamonti, G., Savini, E., Focher, F., Spadari, S. and Ciarrocchi, G. (1992) DNA ligase I gene expression during differentiation and cell proliferation. *Nucleic Acids Res.* **20** (23), 6209-6214.

Montecucco, A., Savini, E., Biamonti, G., Stefanini, M., Focher, F. and Ciarrocchi, G. (1995) Late induction of human DNA ligase I after UV-C irradiation. *Nucleic Acids Res.* **23** (6), 962-966.

Moore, P and Strauss, B.S. (1979) Sites of inhibition of *in vitro* DNA synthesis in carcinogen- and UV-treated phi-X174 DNA. *Nature* **278** 664-666.

Mu, D., Hsu, D.S., and Sancar, A.. (1996) Reaction mechanism of human DNA repair excision nuclease. *J.Biol.Chem.* **271** 8285-8294.

Murray, A and Hunt, T. *The Cell Cycle.*, Oxford University Press, Inc., New York; 1993

Naegli, H. (1995) Mechanisms of DNA damage recognition in mammalian nucleotide excision repair. *FASEB J.* **9** 1043-1050.

Nakane, H., Takeuchi, S., Yuba, S., Saijo,M., Nakatsu, Y., Murai, H., Nakatsuru, Y., Ishikawa, T., Hirota, S., Kitamura, Y., Kato, Y., Tsunoda, Y., Miyauchi, H., Horio, T., Tokunaga, T., Matsunaga, T., Nikaido, O., Nishimune, Y., Okada, Y. and Tanaka, K. (1995) High incidence of

ultraviolet-B-or chemical-carcinogen-induced skin tumours in mice lacking the xeroderma pigmentosum group A gene. *Nature* **377** (6545), 165-168.

Núñez, F., Chipchase, M.D., Clarke, A. and Melton, D.W. (2000) Nucleotide excision repair gene (*ERCC1*) deficiency causes G(2) arrest in hepatocytes and a reduction in liver binucleation: the role of p53 and p21. *FASEB J.* **14** (9), 1073-1082.

O' Donovan, A., Davies, A.A., Moggs, J.G., West, S.C. and Wood, R.D. (1994) XPG endonuclease makes the 3' incision in human DNA nucleotide excision repair. *Nature* **371** (6499), 654-655.

Paques, F. and Haber, J.E. (1997) Two pathways for removal of nonhomologous DNA ends during double-strand break repair in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology* **17** (11), 6765-6771.

Pastink, A., Eeken, J.C. and Lohman, P.H. (2001) Genomic integrity and the repair of double-strand DNA breaks. *Mutat. Res.* **480-481** 37-50.

Pastwa, E. and Blasiak, J. (2003) Non-homologous DNA end joining. *Acta. Biochim. Pol.* **50** (4), 891-908.

Payne, A. and Chu, G. (1994) Xeroderma pigmentosum group-E binding-factor recognizes a broad spectrum of DNA damage. *Mutat.Res.* **310** 89-102.

Pearlman, D.A., and Holbrook, S.R. (1985) Molecular models for DNA damaged by photoreaction. *Science* **227** 1304-1308.

Poole, T. (ed*), The UFAW Handbook on the care and Management of Laboratory Animals 6th edition.* Bath. Longmann Scientific and Technical; 1989.

Ramirez, A., Page, A., Gandarillas, A., Zanet, J., Pibre, S., Vidal, M., Tusell, L., Genesca, A., Whitaker, D.A., Melton, D.W., Jorcano, J.L. (2004) A keratin

K5Cre transgenic line appropriate for tissue-specific or generalized Cre-mediated recombination. *Genesis* **39**(1):52-7.

Ramiro, D., Camenisch, U. and Naegeli, H. (2004) Mechanisms of DNA damage recognition and strand discrimination in human nucleotide excision repair. *DNA Repair* **3** 1409-1423.

Reardon, J.T., Nichols, A.F., Keeney, S., Smith, C.A., Taylor, J.S., Linn, S and Sancar, A (1993) Comparative analysis of binding of human damaged DNA-binding protein (XPE) and *Escherichia coli* damage recognition protein (UvrA) to the major ultraviolet photoproducts-T[c,s]T, T[t,s]T, T[6-4]T, and T. *J.Biol.Chem.* **268** 21301-21308.

Reardon, J.T., Mu, D. and Sancar, A. (1996) Overproduction, purification, and characterization of the XPC subunit of the human DNA repair excision nuclease. *J.Biol.Chem.* **271** 19451-19456.

Reardon, J.T. and Sancar, A. (2002) Molecular anatomy of the human excision nuclease assembled at site of DNA damage. *Mol.Cell.Biol.* **22** 5938-5945.

Reardon, J.T. and Sancar, A. (2003) Recognition and repair of the cyclobutane thymine dimer, a major cause of skin cancers, by human excision nuclease. *Genes Dev.* **17** 2539-2551.

Riedl, T., Hanaoka, F. and Egly, J.M. (2003) The comings and goings of nucleotide excision repair factors on damaged DNA. *EMBO J.* **22** 5293-5303.

Roth, D.B. and Wilson, J.H. (1986) Nonhomologous recombination in mammalian cells: role for short sequence homologies in the joining reaction. *Mol.Cell.Biol.* **6** (4295), 4304.

Rydberg, B. (1978) Bromouracil mutagenesis and mismatch repair in mutator strains of *Escherichia coli. Mutat.Res.* **5** (2), 11-24.

Sancar, A. (1994) Structure and function of DNA photolyase. *Biochemistry.* **33** (1), 2-9.

Sancar, A. (1995) Excision repair in mammalian cells. *J Biol Chem.* **270** (27), 15915-15918.

Sancar, A. (1996) No "End of History" for photolyases. *Science.* **272** (5258), 48-49.

Sancar, A. (1996) DNA excision repair. *Ann.Rev.Biochem.* **65** (43), 81.

Sands, A.T., Abuin, A., Sanchez, A., Conti, C.J. and Bradley, A. (1995) High susceptibility to ultraviolet-induced carcinogenesis in mice lacking XPC. *Nature* **377** 162-165.

Sapabarev, M., Prakash, L. and Prakash, S. (1996) Requirement of mismatch repair genes MSH2 and MSH3 in the RAD1-RAD10 pathway of mitotic recombination *in Saccharomyces cerevisiae*. *Genetics* **142** (3), 727-736.

Sauer, B. (1998). Inducible Gene Targeting in Mice Using the Cre/lox system. *Methods* **14** (4) 381-392.

Setlow, R.B., Swenson, P.A. and Carrier, W.L. (1963) Thymine dimers and inhibition of DNA synthesis by ulatraviolet irradiation of cells. *Science* **142** 1464-1466.

Shiomi, N., Kito, S., Oyama, M., Matsunaga, T., Harada, Y.N., Ikawa, M., Okabe, M. and Shiomi, T. (2004) Identification of the XPG region that causes the onset of Cockayne syndrome by using Xpg mutant mice generated by the cDNA-mediated knock-in method. *Mol. Cell. Biol.* **24** (9), 3712-3719.

Shivji, M.K., Podust, V.N., Hübscher, U. and Wood, R.D. (1995) Nucleotide excision repair DNA synthesis by DNA polymerase in the presence of PCNA, RFC, and RPA. *Biochemistry* **34** 5011-5017.

Sijbers, A.M., van der Spek, P.J., Odijk, H., van den Berg, J., van Duin, M., Westerveld, A., Jaspers, N.G., Bootsma, D. and Hoeijmakers, J.H. (1996) Mutational analysis of the human nucleotide excision repair gene ERCC1. *Nucleic Acids Res.* **24** (17), 3370-3380.

Spivak, G., Itoh, T., Matsunaga, T., Nikaido, O., Hanawalt, P. and Yamaizumi, M. (2002) Ultraviolet-sensitive syndrome cells are defective in transcription-coupled repair of cyclobutane pyrimidine dimers. *DNA Repair* **1** (8), 629-643.

Sugasawa, K., Ng, J.M.Y., Masutani, C., Iwai, S., van der Spek, P.J., Eker, A.P.M., Hanaoka, F., Bootsma, D. and Hoeijmakers, J.H.J. (1998) Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair. *Mol.Cell Biol.* **2** (223), 232.

Sugasawa, K., Okamoto, T., Shimizu, Y., Masutani, C., Iwai, S., and Hanaoka, F. (2001) A multistep damage recognition mechanism for global genomic nucleotide excision repair. *Genes & Dev.* **15** 507-521.

Sugawara, N., Ira, G. and Haber, J.E. (2000) DNA length dependence of the single-strand annealing pathway and the role of Saccharomyces cerevisiae RAD59 in double-strand break repair. *Mol. and Cell. Biol.* **20** (14), 5300-5309.

Sun, D., Urrabaz, R., Nguyen, M., Marty, J., Stringer, S., Cruz, E., Medina-Gundrum, L. and Weitman, S. (2001) Elevated Expression of DNA Ligase I in Human Cancers. *Clinical Cancer Research* **7** 4143-4148.

Sun, X.Z., Harada, Y.N., Zhang, R., Cui, C., Takahashi, S. and Fukui, Y. (2003) A genetic mouse model carrying the nonfunctional xeroderma pigmentosum group G gene. *Congenit. Anom. (Kyoto)* **43** (2), 133-139.

Sutherland, B.M. and Oliver, R. (1976) Culture conditions affect photoreactivating enzyme levels in human fibroblasts. *Biochim Biophys Acta.* **442** (3), 358-367.

Swann, P.F. (1990) Why do O6-alkylguanine and O4-alkylthymine miscode? The relationship between the structure of DNA containing O6-alkylguanine and O4-alkylthymine and the mutagenic properties of these bases. *Mutat Res.* **233** (1-2), 81-94.

Swenson, P.A. and Setlow, R.B. (1966) Effects of ultraviolet radiation on macromolecular synthesis in *Escherichia coli. J. Mol. Biol.* **15** (1), 201-219.

Teo, I., Sedgwick, B., Li, B. and Lindahl, T. (1984) Induction of resistance to alkylating agents in *E. coli*: the ada+ gene product serves both as a regulatory protein and as an enzyme for repair of mutagenic damage. *EMBO J.* **3** (2151), 2157.

Tian, M., Shinkura, R., Shinkura, N. and Alt, F.W. (2004) Growth retardation, early death, and DNA repair defects in mice deficient for the nucleotide excision repair enzyme XPF. *Mol. Cell. Biol.* **24** (3), 1200-1205.

Tomkinson, A.E., Bardwell, A.J., Bardwell, L., Tappe, N.J.and Friedberg, E.C. (1993) Yeast DNA repair and recombination proteins Rad1 and Rad10 constitute a single-stranded-DNA endonuclease. *Nature* **362** (6423), 860-862.

Tornaletti, S. and Hanawalt, P.C. (1999) Effect of DNA lesions on transcription elongation. *Biochimie* **81** 139-146.

van der Horst, G.T., van Steeg, H., Berg, R.J., van Gool, A.J., de Wit, J., Weeda, G., Morreau, H., Beems, R.B., van Kreijl, C.F., de Gruijl, F.R., Bootsma, D., and Hoeijmakers, J.H. (1997) Defective transcription-coupled repair in Cockayne syndrome B mice is associated with skin cancer predisposition. *Cell* **89** 425-435.

van der Horst, G.T., Meira, L., Gorgels, T.G., de Wit, J., Velasco-Miguel, S., Richardson, JA., Kamp, Y., Vreeswijk, M.P., Smit, B., Bootsma, D., Hoeijmakers, J.H. and Friedberg, E.C. (2002) UVB radiation-induced cancer

predisposition in Cockayne syndrome group A (Csa) mutant mice. *DNA Repair (Amst).* **1** (2), 143-157.

van Duin, M., de Wit, J., Odijk, H., Westerveld, A., Yasui, A., Koken, H.M., Hoeijmakers, J.H. and Bootsma, D. (1986) Molecular characterization of the human excision repair gene ERCC-1: cDNA cloning and amino acid homology with the yeast DNA repair gene RAD10. *Cell.* **44** (6), 913-923.

van Duin, M., Koken, M.H., van den Tol, J., ten Dijke, P., Odijk, H., Westerveld, A., Bootsma, D., and Hoeijmakers, J.H. (1987) Genomic characterization of the human DNA excision repair gene ERCC-1. *Nucleic Acids Res.* **15** (22), 9195-9213.

van Duin, M., van den Tol, J., Warmerdam, P., Odijk, H., Meijjer, D., Westerveld, A., Bootsma, D. and Hoeijmakers, J.H.J. (1988) Evolution and mutagenesis of the mammalian excision repair gene ERCC-1. *Nucleic Acids Res.* **16** (12), 5305-5322.

van Duin, M., Vredeveldt, G., Mayne, L.V., Odijk, H., Vermeulen, W., Klein, B., Weeda, G., Hoeijmakers, J.H., Bootsma, D. and Westerveld, A. (1989) The cloned human DNA excision repair gene ERCC-1 fails to correct xeroderma pigmentosum complementation groups A through I. *Mutat.Res.* **217** (2), 83-92.

van Oostrom, C.T., de Vries, A., Verbeek, S.J., van Kreijl, C.F. and van Steeg, H. (1994) Cloning and characterization of the mouse XPAC gene. *Nucleic Acids Res.1994 Jan 11;22(1):11-4.* **22** (1), 11-14.

van Vuuren, A.J., Vermeulen, W., Ma, L., Weeda, G., Appeldoorn, E., Jaspers, N.G., van der Eb, A.J., Bootsma, D., Hoeijmakers, J.H., Humbert, S., Schaeffer, L.and Egly, J.M. (1994) Correction of xeroderma pigmentosum repair defect by basal transcription factor BTF2 (TFIIH). *EMBO J.* **13** (7), 1645-1653.

Venema, J., van Hoffen, A., Karcagi, V., Natarajan, A.T., van Zeeland, A.A. and Mullenders, L.H. (1991) Xeroderma pigmentosum complementation group C cells remove pyrimidine dimers selectively from the transcribed strand of active gene. *Mol.Cell.Biol.* **11** 4128-4134.

Villani, G., Boiteux, S. and Radman, M. (1978) Mechanism of ultraviolet-induced mutagenesis: extent and fidelity of in vitro DNA synthesis on irradiated templates. *Proc. Natl. Acad. Sci. U. S. A.* **75** (7), 3037-3041.

Volker, M., Moné, M.J., Karmakar, P., van Hoffen, A., Schul, W., Vermeulen, W., Hoeijmakers, J.H.J., van Driel, R., van Zeeland, A.A., and Mullenders, L.H.F. (2001) Sequential assembly of the nucleotide excision repair factors in vivo. *Mol.Cell. Biol.* **8** 213-224.

Wakasugi, M. and Sancar, A. (1998) Assembly, subunit composition, and footprint of human DNA repair excision nuclease. *Proc.Natl.Acad.Sci.* **95** 6669-6674.

Wallace, S.S. (1988) AP endonucleases and DNA glycosylases that recognize oxidative damage. *Environ.Mol.Mutagen.* **12** 431-477.

Wang, X.W., Yeh, H., Schaeffer, L., Roy, R., Moncollin, V., Egly, J.M., Wang, Z., Freidberg, E.H., Evans, M.K., Taffe, B.G., Bohr, V.A., Weeda, G., Hoeijmakers, J.H.J., Forrester, K. and Harris, C.C. (1995) p53 modulation of TFIIH-associated nucleotide excision repair activity. *Nat.Genet.* **10** 188-195.

Weeda, G., Donker, I., de Wit, J., Morreau, H., Janssens, R., Vissers, C.J., Nigg, A., van Steeg, H., Bootsma, D., and Hoeijmakers, J.H. (1997) Disruption of mouse ERCC1 results in a novel repair syndrome with growth failure, nuclear abnormalities and senescence. *Curr. Biol.* **7** (6), 427-439.

Weis,L. and Reinberg,D. (1992) Transcription by RNA polymerase II: initiator-directed formation of transcription-competent complexes. *FASEB J.* **6** (14), 3300-3309.

Weis,L.and Reinberg, D. (1997) Accurate positioning of RNA polymerase II on a natural TATA-less promoter is independent of TATA-binding-protein-associated factors and initiator-binding proteins. *Mol. Cell. Biol.* **17** (6), 2973-2984.

Westerveld, A., Hoeijmakers, J.H., van Duin, M., de Wit, J., Odijk, H., Pastink, A., Wood, R.D. and Bootsma, D. (1984) Molecular cloning of a human DNA repair gene. *Nature* **310** (5976), 425-429.

Wijnhoven, S.W.P. and van Steeg, H. (2003) Transgenic and knockout mice for DNA repair functions in carcinogenesis and mutagenesis. *Toxicology* **193** (1-2), 171-187.

Wilkins, R.C. and Lis, J.T. (1998) GAGA factor binding to DNA via a single trinucleotide sequence element. *Nucleic Acids Res.* **26** (11), 2672-2678.

Wilkins, R.C. and Lis, J.T. (1999) DNA distortion and multimerization: novel functions of the glutamine-rich domain of GAGA factor. *J. Mol. Biol.* **15** (285), 515-525.

Wood, R.D. and Burki, H.J. (1982) Repair capability and the cellular age respone for killing and mutation induction after UV. *Mutat.Res.* **95** 505-514.

Wood, R.D. (1996) DNA repair in eukaryotes. *Mutat.Res.* **95** 505-514.

Wu, C.L. and Melton D.W. (1993) Production of a model for Lesch-Nyhan syndrome in hypoxanthine phosphoribosyltransferase-deficient mice. *Nat. Genet.* **3** (3), 235-340.

Wu, X., Wilson, T.E. and Lieber, M.R. (1999) A role for FEN-1 in nonhomologous DNA end joining: The order of strand annealing and nucleolytic processing events. *Proc.Natl.Acad.Sci.U.S.A.* **96** (4), 1303-1308.

Wyse, B.D., Linas, S.L. and Thekkumkara, T.J. (2000) Functional role of a novel cis-acting element (GAGA box) in human type-1 angiotensin II receptor gene transcription. *J. Mol. Endocrinol.* **25** (1), 97-108.

Xu, G. and Goodridge, A.G. (1996) Characterization of a polypyrimidine/polypurine tract in the promoter of the gene for chicken malic enzyme. *J. Biol. Chem.* **271** (27), 16008-16019.

Xu, H., Swoboda, I., Bhalla, P.L., Sijbers, A.M., Zhao, C., Ong, E.K., Hoeijmakers, J.H., and Singh, M.B. (1992) Plant homologue of human excision repair gene ERCC1 points to conservation of DNA repair mechanisms. *Cell* **71** 505-514.

Yasui, A., Takao, M., Oikawa, A., Kiener, A., Walsh, C.T. and Eker, A.P. (1988) Cloning and characterization of a photolyase gene from the cyanobacterium *Anacystis nidulans. Nucleic Acids Res.* **16** (10), 4447-4463.

I Meant to Do My Work Today

Richard LeGallienne

I meant to do my work today -

But a brown bird sang in the apple tree,

And a butterfly flitted across the field,

And all the leaves were calling me.

And the wind sighing over the land

Tossing the grasses to and fro,

And a rainbow held out its shining hand -

So what could I do but laugh and go?